Notes on the Gross-Zagier formula: the case $N = 1$, on singular moduli

Gross and Zagier's 1984 paper [3] studies the $N = 1$ case of the Gross-Zagier formula. In this section I want to study first how this paper fits into the larger formula as a special case and second how to prove it, via both the algebraic and analytic methods (both of which will be needed for the general case).

1. Placement

First: what is a singular modulus? Let

$$j(\tau) = \frac{1}{q} + 744 + 196884q + O(q^2)$$

be the j-invariant. Let K be an imaginary quadratic number field. By the theory of complex multiplication, for any $\tau \in K$ we have $j(\tau) \in H_K$, the Hilbert class field of K, and if τ generates K then $K(j(\tau)) = H_K$ and $j(\tau)$ is an algebraic integer of degree $h(K) = \# \operatorname{Cl}(K) = [H_K : K]$. To each such τ we can associate an elliptic curve E over $\mathbb{C}$ as the quotient of $\mathbb{C}$ by the lattice generated by 1 and τ, with j-invariant $j(E) = j(\tau)$; this elliptic curve has complex multiplication by the order $\mathcal{O}$ generated by τ, and is defined over H_K.

Recalling the setup of the Gross-Zagier formula, a point of $X_0(1)(\mathbb{C})$ corresponds to an elliptic curve E over $\mathbb{C}$, together with an isomorphism $\phi : E \to E$. Last time, we saw that there was a one-to-one correspondence between isogenies $E' \to E''$ of degree N with complex multiplication by $\mathcal{O}_K$ and ideals I of $\mathcal{O}_K$ with $\mathcal{O}_K/I \simeq \mathbb{Z}/N\mathbb{Z}$; setting $N = 1$, the only such ideal I is $I = \mathcal{O}_K$, and so there is a unique E' with complex multiplication by $\mathcal{O}_K$, which is defined over H_K; this is precisely the elliptic curve defined above.

For the more general Gross-Zagier formula, we would then take an elliptic curve over $\mathbb{Q}$ of discriminant N and an imaginary quadratic number field K of discriminant d_K, let x_K be a unique point of $X_0(N)(H_K)$ with complex multiplication, as above. But for $N = 1$ no such elliptic curves exist! So interpreted literally the $N = 1$ case is trivial.

The proof, though, is genuinely (an extension of) a special case. Recall that the proof of the Gross-Zagier formula involves computing both sides as pairings with some cusp forms F and G: in particular the Fourier coefficients of G are given by the sum of local height pairings $\langle x_K, T_n x_K \rangle$ for x_K a Heegner point on $X_0(N)$, i.e. a point with complex multiplication by K. At the finite places the height pairings reduce to questions about the endomorphisms of elliptic curves over quotients of extensions; these are nontrivial at the supersingular primes, where we can do an explicit calculation in the corresponding quaternion algebra. In the infinite places the pairing is given by the solution to a certain differential equation, which can be solved explicitly, after some ideal-counting. But in this case the global pairing is 0 everywhere since $X_0(1) \simeq \mathbb{P}^1$ and therefore every degree 0 line bundle is trivial, so the contributions from the finite and infinite places are equal up to a sign. Thus in this case although the global result is trivial, we can still make the local computations in both the archimedean and nonarchimedean case; this gives two proofs of the same equality, algebraic and analytic, which can be massaged into the following form.

Let d_1 and d_2 be relatively prime fundamental discriminants corresponding to orders in which there are w_1 and w_2 roots of unity respectively, and let j be the j-invariant. Define

$$J(d_1, d_2) = \prod_{\tau_1, \tau_2} (j(\tau_1) - j(\tau_2))^{\frac{4}{w_1 w_2}},$$

where the product is over equivalence classes of algebraic numbers τ_1, τ_2 satisfying $a_i\tau_i^2 + b_i\tau_i + c = 0$ with $\mathrm{disc}(\tau_i) = b_i^2 - 4a_ic_i = d_i$ for $i = 1, 2$.

Let p be a prime number, and write $\left(\frac{a}{p}\right)$ for the Legendre symbol. Since d_1 and d_2 are relatively prime, p divides at most one of them, so at least one of $\left(\frac{d_1}{p}\right)$ and $\left(\frac{d_2}{p}\right)$ is nonzero; if p divides neither, so that $\left(\frac{d_1 d_2}{p}\right) = \left(\frac{d_1}{p}\right)\left(\frac{d_2}{p}\right) \neq 0$, then if $\left(\frac{d_1 d_2}{p}\right) = 1$ then $\left(\frac{d_1}{p}\right) = \left(\frac{d_2}{p}\right)$. Therefore for any p such that $\left(\frac{d_1 d_2}{p}\right) \neq -1$ we can define $\epsilon(p)$ to be whichever of $\left(\frac{d_i}{p}\right)$ is nonzero, since at least one is nonzero and if both are then they are equal. We can extend ϵ to all natural numbers by multiplicativity.

Theorem 1.1. *With notation as above,*

$$J(d_1, d_2)^2 = \pm \prod_{\substack{n, n' \geq 1 \\ x \in \mathbb{Z} \\ x^2 + 4nn' = d_1 d_2}} n^{\epsilon(n')}.$$

This is well-defined: since n' divides $\frac{d_1 d_2 - x^2}{4}$, for every prime p dividing n' we have $d_1 d_2 - x^2 = 0 \pmod{p}$ and so $\left(\frac{d_1 d_2}{p}\right) \neq -1$. It is easy to understand from this for example why the integers $J(d_1, d_2)^2$ have such small prime factors: for p dividing $J(d_1, d_2)^2$, we must have p dividing some n which satisfies $nn' \leq \frac{d_1 d_2}{4}$, and so $p \leq \frac{d_1 d_2}{4}$. This is remarkable since $J(d_1, d_2)^2$ may be very large: for example, for $d_1 = -67$ and $d_2 = -163$, so that the class numbers of the corresponding orders are both 1 so that J is a single factor

$$\begin{aligned} J(-67, -163) &= j\left(\frac{1 + \sqrt{-67}}{2}\right) - j\left(\frac{1 + \sqrt{-163}}{2}\right) \\ &= -147197952000 + 262537412640768000 \\ &= 262537265442816000 \end{aligned}$$

as is easily computed from the q-expansion of the j-invariant, as we know that since the class numbers are 1 these are both integers. The prime factorization of this number is

$$2^{15} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 139 \cdot 331$$

(compare the bound $\frac{67 \cdot 163}{4} = \frac{10921}{4} = 2730.25$); the proportion of integers N with largest prime factor at most $n^{\log_{262537265442816000} 331} = n^{.144658\ldots}$ is approximately $1.156 \cdot 10^{-6}$, via the Dickman–de Bruijn rho function, so this is indeed quite unusually smooth. (If we were to instead use the bound 2730.25 rather than the true answer 331, the proportion of integers with the corresponding bound would be approximately $2.926 \cdot 10^{-4}$.)

But in fact, we can understand the smoothness of $J(d_1, d_2)^2$ through much less work than it takes to prove Theorem 1.1. We can understand $J(d_1, d_2)$ as the norm of $j(\tau_1) - j(\tau_2)$ for some τ_i of discriminant d_i; for simplicity let's think about the case where the $j(\tau_i)$ are themselves integers, which occurs when both number fields $\mathbb{Q}(\sqrt{d_i})$ have class number 1. What condition can we put on primes dividing $j(\tau_1) - j(\tau_2)$?

Well, if E_1 and E_2 are the corresponding elliptic curves of discriminant d_1 and d_2 respectively, then p divides $j(\tau_1) - j(\tau_2) = j(E_1) - j(E_2)$ only if E_1 is isomorphic to E_2 after reduction modulo p. Since each E_i has complex multiplication by an order $\mathcal{O}_i$ in $\mathbb{Q}(\sqrt{d_i})$, if $E_p \simeq E_1 \simeq E_2 \pmod p$ is the modulo p elliptic curve it carries an action of both $\mathcal{O}_1$ and $\mathcal{O}_2$, i.e. we have embeddings $\mathcal{O}_1 \hookrightarrow \mathrm{End}(E_p) \hookleftarrow \mathcal{O}_2$.

Recall that the endomorphism ring of an elliptic curve defined over $\mathbb{F}_p$ is either an order in an imaginary quadratic field (the ordinary case) or an order in the quaternion algebra over $\mathbb{Q}$ ramified at ∞ and at p (the supersingular case). If $\mathrm{End}(E_p)$ is an order in an imaginary quadratic field K of discriminant d_K, then since it contains two orders $\mathcal{O}_1$ and $\mathcal{O}_2$ of imaginary quadratic fields we must have $d_K | d_1$ and $d_K | d_2$; since we have assumed that $\gcd(d_1, d_2) = 1$, this is impossible unless $d_K = 1$, which does not hold for any imaginary quadratic field. Therefore we must be in the supersingular case: $\mathrm{End}(E_p)$ is an order in the quaternion algebra ramified at ∞ and at p.

Suppose that p splits in $\mathcal{O}_1$. Then $\mathcal{O}_1 \otimes \mathbb{Q}_p$ is isomorphic to the direct product $\mathbb{Q}_p \times \mathbb{Q}_p$; since $\mathbb{Q}_p$ is torsion-free it is flat and so we again have an injection $\mathcal{O}_1 \otimes \mathbb{Q}_p \hookrightarrow \mathrm{End}(E_p) \otimes \mathbb{Q}_p$, which by assumption is a division ring. But $\mathcal{O}_1 \otimes \mathbb{Q}_p \simeq \mathbb{Q}_p \times \mathbb{Q}_p$ contains zero divisors, e.g. $(0, 1)$ and $(1, 0)$, while $\mathrm{End}(E_p) \otimes \mathbb{Q}_p$ is a division ring and therefore contains no zero divisors; therefore no such injection can exist, and so p is inert or ramified in $\mathcal{O}_1$, and by the same argument in $\mathcal{O}_2$. Suppose for simplicity that $p \nmid d_1 d_2$, so p is inert in both $\mathcal{O}_1$ and $\mathcal{O}_2$; then $\left(\frac{d_1}{p}\right) = \left(\frac{d_2}{p}\right) = -1$, and so $\left(\frac{d_1 d_2}{p}\right) = 1$. Thus there exists some integer x such that $d_1 d_2 \equiv x^2 \pmod p$ and therefore $p | d_1 d_2 - x^2$. Indeed, $\mathcal{O}_1$ and $\mathcal{O}_2$ together generate an order $\mathcal{O} \subseteq \mathrm{End}(E_p)$, corresponding to an action of $\mathbb{Q}(\sqrt{d_1 d_2})$, whose discriminant must be of the form $x^2 - d_1 d_2$ and must be negative; therefore we have $p | d_1 d_2 - x^2 > 0$. Other than a slight variation at $p = 2$, this is the claimed smoothness condition.

What Theorem 1.1 gives beyond this is a precise formula for not only which primes divide $J(d_1, d_2)^2$, but also to which powers.

2. Algebraic proof

First, let's go through the steps of the proof without proving the requisite lemmas so as to understand the structure, and come back to the proofs at the end.[1]

Fix an imaginary quadratic number field $K = \mathbb{Q}(\tau)$, and let $\mathcal{O} = \mathbb{Z}[\tau] \subseteq \mathcal{O}_K$ be an order of $\mathcal{O}_K$; define d to be the discriminant of this order, i.e. if τ satisfies the equation $a\tau^2 + b\tau + c = 0$ for a, b, c integers, then $d = b^2 - 4ac$. For simplicity suppose that $d \not\equiv 1 \pmod 4$, so that $\mathcal{O} = \mathbb{Z}[\tau] = \mathcal{O}_K$, and that $d < -4$; then $K(j(\tau)) = H_K$, and the number of

[1] Or just skip them entirely actually; in theory I'll come back and actually fill them in at some point.

roots of unity in $\mathcal{O}$ is precisely 2. Define

$$\alpha(\tau, d_2) = \prod_{\substack{\tau_2 \\ \mathrm{disc}(\tau_2)=d_2}} (j(\tau) - j(\tau_2))^{\frac{4}{2w_2}}$$

where τ_2 ranges over classes of algebraic numbers τ_2 such that $\mathbb{Z}[\tau_2]$ is an order of discriminant $d_2 < 0$ and w_2 is the number of roots of unity in $\mathbb{Z}[\tau_2]$, where τ_2 and τ_2' are considered equivalent if the corresponding quadratic forms are equivalent, i.e. if $a\tau_2^2 + b\tau_2 + c = 0$ and $a'\tau_2'^2 + b'\tau_2 + c' = 0$ then there exist integers α, β such that $ax^2 + bxy + cy^2 = a'(\alpha x + \beta y)^2 + b'(\alpha x + \beta) + c'$ for all x, y. Thus

$$J(d, d_2) = \prod_{\substack{\tau \\ \mathrm{disc}(\tau)=d}} \alpha(\tau, d_2)$$

with the same conventions as above.

The set of τ_2 with discriminants equal to d_2 has cardinality equal to the class number $h(d_2)$ of $\mathbb{Z}[\sqrt{-d_2}]$; again for simplicity we assume that $d_2 \not\equiv 1 \pmod 4$, so that this is the ring of integers of $L := \mathbb{Q}(\sqrt{-d_2})$, and that $d_2 < -4$ so that $w_2 = 2$; we'll also assume that d and d_2 are relatively prime. Therefore by the theory of complex multiplication each $j(\tau_2)$ is an algebraic integer over $\mathbb{Q}$ of degree $h(d_2)$, and indeed all of the $j(\tau_2)$ are Galois conjugates so that

$$G(x) := \prod_{\substack{\tau_2 \\ \mathrm{disc}(\tau_2)=d_2}} (x - j(\tau_2))$$

is defined over $\mathbb{Q}$; therefore

$$\alpha(\tau, d_2) = G(j(\tau))$$

is in $\mathbb{Q}(j(\tau)) \subseteq H_K$. Our goal will then be to evaluate $\alpha(\tau, d_2)$ up to a unit by computing its valuation $v(\alpha)$ at each finite place v of H_K; this will fully determine the ideal (α) generated by α, and therefore once we do this for each τ of discriminant d we can fully determine the ideal generated by $J(d, d_2)$ in the ring of integers of H_K. Since H_K is the Hilbert class field of K and $J(d, d_2)$ is an integer under our assumptions, this determines the prime factorization of $J(d, d_2)$ and therefore fixes its value up to a unit.

Let H_v be the completion of H_K at a finite place v lying over a prime p and let H_v^{unr} be the maximal unramified extension of H_v, so that the residue field of H_v^{unr} is the union of $\mathbb{F}_{q^r}$ over all r, i.e. $\overline{\mathbb{F}}_q$, where q is the order of the residue field of H_v, or equivalently the norm of the prime ideal $\mathfrak{p}$ associated to v. We can think of H_v^{unr} as the ring of Witt vectors of $\overline{\mathbb{F}}_q$. Fix some γ satisfying an equation $a\gamma^2 + b\gamma + c = 0$ of discriminant d_2 (recall that this is the same requirement that the various τ_2 must satisfy, so this is equivalent to fixing one of the τ_2), and let $W = \mathcal{O}_{\widehat{H}_v^{\mathrm{unr}}}[\gamma]$ where $\widehat{H}_v^{\mathrm{unr}}$ is the completion of H_v^{unr}. To evaluate $\alpha(\tau, d_2)$, we work one factor at a time in this larger setting: fix some elliptic curve E over W with complex multiplication by $\mathcal{O} = \mathcal{O}_K = \mathbb{Z}[\tau]$ and $j(E) = j(\tau)$, and for each τ_2 let E' be an elliptic curve over W with complex multiplication by $\mathbb{Z}[\gamma]$ and $j(E') = j(\tau_2)$. (These exist be a result of Serre and Tate [4].) We now need our first lemma.

Lemma 2.1. *Let W be a complete discrete valuation ring, with fraction field of characteristic 0 and residue field of characteristic $p > 0$, and let π be a uniformizer with valuation v*

normalized so that $v(\pi) = 1$. Let E, E' be elliptic curves over W with good reduction modulo π and distinct j-invariants, and let $i(E, E', n)$ be half the number of isomorphisms $E \to E'$ modulo π^n. Then

$$v(j(E) - j(E')) = \sum_{n=1}^{\infty} i(E, E', n).$$

This is proven by choosing a model for each curve and using the explicit formula for j in terms of these models and similarly explicitly counting isomorphisms, and checking that in all cases these are equal.

It is worth remarking first that the number of isomorphisms $E \to E'$ is always even: given an isomorphism f, we can define $\tilde{f} : E \to E'$ sending $P \mapsto f(-P) = -f(P)$, which satisfies $\tilde{f}(P+Q) = f(-P-Q) = -f(P)-f(Q) = \tilde{f}(P)+\tilde{f}(Q)$ and so is also a homomorphism, and is clearly also a bijection and distinct from f. There is an injection from the set of isomorphisms $E \to E'$ to the set of automorphisms of E, sending a pair of distinct isomorphisms $f, g : E \to E'$ (with g not equal to $\tilde{f}$ unless these are the only two isomorphisms) to $f^{-1}g$ and $g^{-1}f$, and there are finitely many automorphisms of E (0, 2, 4, 6, 12, or 24, with the last two impossible in characteristic $p > 3$).

Applying this lemma, we can reduce our problem to one of counting isomorphisms $E \to E'$ modulo π^n, with E and E' as above. Explicitly,

$$v(\alpha) = \frac{1}{e} \sum_{\substack{\tau_2 \\ \operatorname{disc}(\tau_2)=d_2}} v(j(\tau) - j(\tau_2)) = \frac{1}{e} \sum_{\substack{E' \\ \operatorname{disc}(j(E'))=d_2}} \sum_{n=1}^{\infty} i(E, E', n)$$

where the first sum is taken over E' with complex multiplication by γ and e is the ramification index of W, by which we divide to normalize the difference between the valuations of W and of $\widehat{H}_v^{\mathrm{unr}}$; we will generally assume that this ramification is 1 for simplicity.We have a distinguished endomorphism of E' given by multiplication by γ, which we also write as γ; if $f : E \to E'$ is an isomorphism over W/π^n, define the endomorphism $\gamma_f = f^{-1} \circ \gamma \circ f$ of E. By the cyclicity of the trace and commutativity of the norm, this has trace $\operatorname{Tr}(\gamma_f) = \operatorname{Tr}(ff^{-1}\gamma) = \operatorname{Tr}(\gamma)$ and $\mathrm{N}(\gamma_f) = \mathrm{N}(ff^{-1}\gamma) = \mathrm{N}(\gamma)$ in $\operatorname{End}_{W/\pi^n}(E)$. Any endomorphism of E induces an endomorphism of its tangent space at the identity, which since E is a curve is just the one-dimensional module W/π^n over itself; the space of linear endomorphisms of W/π^n is the space of 1×1 matrices with entries in W/π^n, i.e. W/π^n itself, and so the associated endomorphism is simply an element of W/π^n. Since this ring is commutative, the endomorphism associated to $\gamma_f = f^{-1}\gamma f$ is simply γ.

Lemma 2.2. *Let $\phi : E \to E$ be an endomorphism over W/π^n, with notation as above, such that $\operatorname{Tr}(\phi) = \operatorname{Tr}(\gamma)$, $\mathrm{N}(\phi) = \mathrm{N}(\gamma)$, and ϕ induces multiplication by γ on the tangent space to E at the identity. Then there exists a unique elliptic curve E' over W with complex multiplication by $\mathbb{Z}[\gamma]$ and good reduction modulo π and an isomorphism $f : E \to E'$ over W/π^n, unique up to W-automorphisms of E', such that $\phi = \gamma_f$.*

Letting S_n be the set of endomorphisms ϕ of E over W/π^n satisfying the hypotheses of the lemma, we have

$$\sum_{\substack{E' \\ \operatorname{disc}(j(E'))=d_2}} i(E, E', n) = |S_n|,$$

with the sum as above over E' with complex multiplication by $Z[\gamma]$, since the W-automorphisms of each E' are precisely the number of roots of unity in the corresponding $\mathbb{Z}[\tau_2]$, of which by assumption there are 2. Thus

$$v(\alpha) = \sum_{n=1}^{\infty} |S_n|.$$

Suppose that E has ordinary reduction modulo π. Then $\mathrm{End}_{W/\pi^n}(E) \simeq \mathcal{O} = \mathbb{Z}[\tau]$ for all n, and since we have assumed that d and d_2 are relatively prime $\mathcal{O}$ contains no elements of discriminant d_2, and therefore no elements with trace and norm equal to those of γ. Therefore S_n is empty for every n and so in this case $v(\alpha) = 0$.

Thus the only remaining case is when E has supersingular reduction modulo π, so that $\mathrm{End}_{W/\pi}(E)$ is isomorphic to a maximal order $\mathcal{O}_B$ in the quaternion algebra B over $\mathbb{Q}$ ramified at p and ∞. In this case, $\mathbb{Q}(j(\tau))$ embeds uniquely into $\mathbb{Q}_p$ [2]. Since $H_K = K(j(\tau))$ is a quadratic extension of $\mathbb{Q}(j(\tau))$, there are two ways of extending this embedding to H_K, corresponding to v and some other place v_1 over p; these are Galois conjugates, and so there exists some $\sigma \in \mathrm{Gal}(H_K/K) \simeq \mathrm{Cl}(K)$ such that $v_1(\sigma(\beta)) = v(\beta)$ for all nonzero β in H_K. Let $\mathfrak{a}$ be an ideal corresponding to σ under the Artin isomorphism $\mathrm{Gal}(H_K/K) \simeq \mathrm{Cl}(K)$.

Lemma 2.3. *Let m_n be the number of solutions $(x, \mathfrak{b})$ to the equation $x^2 + 4p^{2n-1}\,\mathrm{N}(\mathfrak{b}) = dd_2$, where $\mathfrak{b}$ is an ideal of $\mathcal{O} = \mathbb{Z}[\tau]$ in the class of $\mathfrak{a}^2$, x is an integer, and any solution $(x, \mathfrak{b})$ with x divisible by d is counted twice.*

1) If $p \nmid dd_2$, then $|S_n| = m_n$ for all $n \geq 1$.

2) If $p | dd_2$, then $|S_1| = m_1$ and S_n is empty for all $n \geq 2$.

If we write $r_{\mathfrak{a}}(k)$ for the number of ideals of $\mathcal{O}$ in the class of $\mathfrak{a}$ of norm k (which is 0 for k not a positive integer), then

$$m_n = \frac{1}{2} \sum_{x \in \mathbb{Z}} r_{\mathfrak{a}^2}\left(\frac{dd_2 - x^2}{4p^{2n-1}}\right).$$

Thus all in all we have

$$v(\alpha) = \frac{1}{2} \sum_{n \geq 1} \sum_{x \in \mathbb{Z}} r_{\mathfrak{a}^2}\left(\frac{dd_2 - x^2}{4p^{2n-1}}\right).$$

Now, $J(d, d_2)$ is the norm

$$J(d, d_2) = \mathrm{N}_{H_K/K}(\alpha) = \prod_{\sigma \in \mathrm{Gal}(H_K/K)} \sigma(\alpha),$$

and σ acts on this formula by permuting the $\mathfrak{a} \in \mathrm{Cl}(K) \simeq \mathrm{Gal}(H_K/K)$. Write $r_{\mathrm{sq}}(k)$ for the set of ideals of $\mathcal{O}$ of norm k whose class is a square in the class group, i.e. is in the image of $\mathfrak{a} \mapsto \mathfrak{a}^2$; for $\mathrm{Cl}(K)$ odd, as happens for example when $-d$ is a prime congruent to 3 modulo 4, this is the entire class group. Then

$$v(J(d, d_2)) = \frac{1}{2} \sum_{n \geq 1} \sum_{x \in \mathbb{Z}} r_{\mathrm{sq}}\left(\frac{dd_2 - x^2}{4p^{2n-1}}\right).$$

Assume that we are in such a situation with $\mathrm{Cl}(K)$ odd,[2] and write $r(k)$ for the number of ideals of $\mathcal{O}_K$ of norm k. Let

$$\zeta_K(s) = \sum_{k=1}^{\infty} r(k) k^{-s}$$

be the Dedekind zeta function of K. We have $\zeta_K(s) = \zeta(s)L(\chi, s)$, where $\zeta(s)$ is the Riemann zeta function and χ is the real quadratic character such that $\chi(p) = 1$ if p splits in $\mathcal{O}$ and $\chi(p) = -1$ otherwise, i.e. $\chi(p) = \left(\frac{d}{p}\right)$. Therefore r is given by Dirichlet convolution

$$r(k) = \sum_{T|k} \chi(T),$$

and so

$$v(J(d, d_2)) = \frac{1}{2} \sum_{n\geq 1} \sum_{x\in\mathbb{Z}} \sum_{Tp^{2n-1} | \frac{dd_2 - x^2}{4}} \chi(T).$$

On the other hand, consider the right-hand side of the formula in Theorem 1.1

$$\prod_{\substack{n,n'\geq 1 \\ x\in\mathbb{Z} \\ x^2+4nn'=dd_2}} n^{\epsilon(n')}.$$

If we let

$$F(m) = \prod_{\substack{n,n'\geq 1 \\ nn'=m}} n^{\epsilon(n')}$$

then this is

$$\prod_{x^2<d_1d_2} F\left(\frac{d_1d_2 - x^2}{4}\right).$$

Write $m = \lambda_1^{2a_1+1} \cdots \lambda_t^{2a_t+1} \cdot \ell_1^{2b_1} \cdots \ell_s^{2b_s} \cdot q_1^{c_1} \cdots q_r^{c_r}$ where $\epsilon(\lambda_i) = \epsilon(\ell_i) = -1$ and $\epsilon(q_i) = 1$ (which is well-defined, since when $m = \frac{d_1d_2 - x^2}{4}$ we always have $\left(\frac{d_1d_2}{q}\right) \neq -1$ for any $q|m$). Clearly if $p \nmid m$ then $p \nmid F(m)$, so suppose that p divides m. We have

$$v(F(m)) = \sum_{\substack{n,n'\geq 1 \\ nn'=m}} \epsilon(n')v(n) = \sum_{k\geq 1} \sum_{\substack{n,n'\geq 1 \\ nn'=m \\ v(n)=k}} \epsilon(n')k.$$

Note that $\epsilon\left(\frac{d_1d_2 - x^2}{4}\right) = -1$, so t must be odd.[3] For every divisor n' of m, we have $n' = \lambda_1^{\alpha_1} \cdots \lambda_t^{\alpha_t} \cdot \ell_1^{\beta_1} \cdots \ell_s^{\beta_s} \cdot q_1^{\gamma_1} \cdots q_r^{\gamma_r}$ for $\alpha_i \leq 2a_i + 1$, $\beta_i \leq 2b_i$, and $\gamma_i \leq c_i$. Thus this is

$$\sum_{k\geq 1} k \sum_{\{\alpha_i, \beta_i, \gamma_i\}} \epsilon\left(\lambda_1^{\alpha_1} \cdots \lambda_t^{\alpha_t} \cdot \ell_1^{\beta_1} \cdots \ell_s^{\beta_s} \cdot q_1^{\gamma_1} \cdots q_r^{\gamma_r}\right)$$

[2]Gross and Zagier specialize to the case where $-d$ and $-d_2$ are primes, in which case this holds; the general case follows from genus theory, which I'll try and put in at some point.

[3]Gross and Zagier state this as if it should be self-evident, but I don't immediately see why.

where the sum is taken over the $\alpha_i, \beta_i, \gamma_i$ satisfying the bounds above except for whichever is the exponent of p, i.e. for example if $p = \ell_j$ then we sum over all α_i, all γ_i, and all β_i except for β_j, which is fixed to be $v(m) - k$. Since we know the value of ϵ on each of these primes, this is

$$\sum_{k\geq 1} k \sum_{\{\alpha_i,\beta_i,\gamma_i\}} (-1)^{\alpha_1+\cdots+\alpha_t+\beta_1+\cdots+\beta_s}$$

where the sum is as above. If p is one of the q_i, since the sum is independent of the c_i it is 0 by symmetry. If $p = \ell_j$, then this is

$$\sum_{k\geq 1} k \sum_{\alpha_1=0}^{2a_1+1} \cdots \sum_{\alpha_t=0}^{2a_t+1} \sum_{\beta_1=0}^{2b_1} \cdots \sum_{\beta_{j-1}=0}^{2b_{j-1}} \sum_{\beta_{j+1}=0}^{2b_{j+1}} \cdots \sum_{\beta_s=0}^{2b_s} \sum_{\gamma_1=0}^{c_1} \cdots \sum_{\gamma_r=0}^{c_r} (-1)^{\alpha_1+\cdots+\alpha_t+b_1+\cdots+b_{j-1}+2b_j-k+b_{j+1}+\cdots+b_s}$$

which simplifies to

$$\sum_{k\geq 1} k(-1)^k \left(\prod_{i=1}^{t} \frac{(-1)^{2a_i+2}-1}{-1-1}\right) \left(\prod_{i'\neq j} \frac{(-1)^{2b_{i'}+1}-1}{-1-1}\right) (c_1+1)\cdots(c_r+1) = 0$$

since $t \neq 0$. The same argument shows that this is 0 if $p = \lambda_j$, unless $t = 1$ so that the product over $i \neq j$ is empty; thus the only case in which $v(F(m))$ can be nonzero is if $m = p^{2a+1} \cdot \ell_1^{2b_1} \cdots \ell_s^{2b_s} \cdot q_1^{c_1} \cdots q_r^{c_r}$ with $\epsilon(p) = \epsilon(\ell_i) = -1$ and $\epsilon(q_i) = 1$.

In this case, the above formula becomes

$$v(F(m)) = \sum_{1\leq k\leq 2a+1} k(-1)^{2a+1-k}(c_1+1)\cdots(c_r+1) = (a+1)(c_1+1)\cdots(c_r+1).$$

Since this is the only prime p for which this is nonzero, we conclude that if $m = p^{2a+1} \cdot \ell_1^{2b_1} \cdots \ell_s^{2b_s} \cdot q_1^{c_1} \cdots q_r^{c_r}$ as above then

$$F(m) = p^{(a+1)(c_1+1)\cdots(c_r+1)}$$

and $F(m) = 1$ otherwise. Therefore

$$v\left(\prod_{x^2<d_1d_2} F\left(\frac{d_1d_2-x^2}{4}\right)\right)$$

can be thought of as the sum over all x such that $\frac{d_1d_2-x^2}{4}$ is of this form, with p the unique prime with $\epsilon(p) = -1$ dividing $\frac{d_1d_2-x^2}{4}$ an odd number of times, of $(a+1)(c_1+1)\cdots(c_r+1)$ where $2a+1$ is the exponent of p and the c_i are the exponents of the primes dividing $\frac{d_1d_2-x^2}{4}$ such that $\epsilon(q) = 1$.

Recall our formula from above

$$v(J(d_1,d_2)) = \frac{1}{2}\sum_{n\geq 1}\sum_{x\in\mathbb{Z}} \sum_{Tp^{2n-1}|\frac{d_1d_2-x^2}{4}} \chi(T),$$

where χ is the real quadratic character of $\mathcal{O} = \mathcal{O}_1$. Assuming for simplicity that $p \nmid d_1$, note that $\chi(T) = \epsilon(T)$. If $\frac{d_1 d_2 - x^2}{4} = \lambda_1^{2a_1+1} \cdots \lambda_t^{2a_t+1} \cdot \ell_1^{2b_1} \cdots \ell_s^{2b_s} \cdot q_1^{c_1} \cdots q_r^{c_r}$ as above, then $T | \frac{d_1 d_2 - x^2}{4p^{2n-1}}$ is given by $T = \lambda_1^{\alpha_1} \cdots \lambda_t^{\alpha_t} \ell_1^{\beta_1} \cdots \ell_s^{\beta_s} \cdot q_1^{\gamma_1} \cdots q_r^{\gamma_r}$ for $\alpha_i \le 2a_i + 1$, $\beta_i \le 2b_i$, and $\gamma_i \le c_i$; a calculation similar to that above shows that the innermost sum is $(c_1+1) \cdots (c_r+1)$ if $t = 1$ and 0 otherwise, and summing over $n \le a_1 + 1$ gives that this is the sum over x such that $\frac{d_1 d_2 - x^2}{4}$ is of this form $p^{2a+1} \ell_1^{2b_1} \cdots \ell_s^{2b_s} \cdot q_1^{c_1} \cdots q_r^{c_r}$ of $\frac{1}{2}(a_1+1)(c_1+1) \cdots (c_r+1)$. Therefore combining this with the above we conclude that

$$2v(J(d_1, d_2)) = v \left(\prod_{\substack{n, n' \ge 1 \\ x \in \mathbb{Z} \\ x^2 + 4nn' = dd_2}} n^{\epsilon(n')} \right)$$

and therefore by doing this at every prime p we conclude that, up to a unit of $\mathbb{Z}$, the expressions whose valuations we are studying must be equal; this is Theorem 1.1.

3. Analytic proof

Again, this is a computation. We take logarithms and compute each $\log |j(\tau_1) - j(\tau_2)|$ as an infinite sum; then we find that we can combine these in such a way as to get a finite sum, which upon exponentiation gives the right-hand side of Theorem 1.1.

Write $\tau_m = u_m + v_m i$ for $m = 1, 2$, and define the functions

$$Q_{s-1}(t) = \int_0^\infty (t + \sqrt{t^2 - 1} \cosh \xi)^{-s} \, d\xi,$$

the hyperbolic distance

$$d(\tau_1, \tau_2) = \cosh^{-1} \left(1 + \frac{(u_1 - u_2)^2 + (v_1 - v_2)^2}{2 v_1 v_2} \right),$$

$$g_s(\tau_1, \tau_2) = -2 Q_{s-1}(\cosh d(\tau_1, \tau_2)) = -2 Q_{s-1} \left(\frac{(u_1 - u_2)^2 + v_1^2 + v_2^2}{2 v_1 v_2} \right),$$

and

$$G_s(\tau_1, \tau_2) = \sum_{\gamma \in \Gamma} g_s(\tau_1, \gamma \tau_2),$$

where $\Gamma = \mathrm{PSL}(2, \mathbb{Z})$. Since $d(\gamma\tau_1, \gamma\tau_2) = d(\tau_1, \tau_2)$ for any γ, we have

$$G_s(\delta\tau_1, \tau_2) = -2 \sum_{\gamma \in \Gamma} Q_{s-1}(\cosh d(\delta\tau_1, \gamma\tau_2)) = -2 \sum_{\gamma \in \Gamma} Q_{s-1}(\cosh d(\tau_1, \delta^{-1}\gamma\tau_2)) = G_s(\tau_1, \tau_2)$$

and

$$G_s(\tau_1, \delta\tau_2) = \sum_{\gamma \in \Gamma} g_s(\tau_1, \gamma\delta\tau_2) = G_s(\tau_1, \tau_2)$$

for any $\delta \in \Gamma$ by relabeling the sum by $\delta^{-1}\gamma \mapsto \gamma$ and $\gamma\delta \mapsto \gamma$. This is real and analytic away from $\tau_1 = \tau_2$. Finally

$$E(\tau, s) = \frac{1}{2} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1}} \frac{v^s}{|c\tau + d|^{2s}}$$

is the Eisenstein series, where v is the imaginary part of τ, and

$$\varphi(s) = \frac{\sqrt{\pi}\Gamma\left(s - \frac{1}{2}\right)\zeta(2s-1)}{\Gamma(s)\zeta(2s)}.$$

Proposition 3.1. *Let τ_1, τ_2 be points in the upper half-plane in distinct orbits under the action of Γ. Then*

$$\log|j(\tau_1) - j(\tau_2)|^2 = \lim_{s \to 1}\left(G_s(\tau_1, \tau_2) + 4\pi(E(\tau_1, s) + E(\tau_2, s) - \varphi(s))\right) - 24.$$

Proof sketch. First observe that the limit on the right-hand side is a sum of four terms each of which has a pole at $s = 1$; by taking Fourier expansions, it is possible to see that each of these poles is simple, with residues -12, 12, 12, and -12 respectively, so in the limit the poles cancel and so the limit exists.

Next, fix τ_2, and consider the differential operator

$$\Delta_1 = v_1^2\left(\frac{\partial^2}{\partial u_1^2} + \frac{\partial^2}{\partial v_2^2}\right).$$

We can compute

$$\begin{aligned}
\Delta_1 G_s(\tau_1, \tau_2) &= \sum_{\gamma \in \Gamma} \Delta_1 g_s(\tau_1, \gamma\tau_2) \\
&= -2\sum_{\gamma \in \Gamma} \Delta_1 Q_{s-1}\left(\frac{(u_1 - \gamma u_2)^2 + v_1^2 + (\gamma v_2)^2}{2v_1\gamma v_2}\right) \\
&= -2\sum_{\gamma \in \Gamma}\left(\frac{((u_1 - \gamma u_2)^2 + (v_1 - \gamma v_2)^2)((u_1 - \gamma u_2)^2 + (v_1 + \gamma v_2)^2)}{4v_1^2(\gamma v_2)^2} Q''_{s-1}\right. \\
&\qquad \left. + \frac{(u_1 - \gamma u_2)^2 + v_1^2 + (\gamma v_2)^2}{v_1\gamma v_2} Q'_{s-1}\right)\left(\frac{(u_1 - \gamma u_2)^2 + v_1^2 + \gamma v_2^2}{2v_1 v_2}\right)
\end{aligned}$$

where γu_2 indicates the real part of $\gamma\tau_2$ and similarly for γv_2. Setting $t = \frac{(u_1 - \gamma u_2)^2 + v_1^2 + (\gamma v_2)^2}{2v_1\gamma v_2}$, this simplifies to

$$\Delta_1 G_s(\tau_1, \tau_2) = -2\sum_{\gamma \in \Gamma}\left((t^2 - 1)Q''_{s-1}(t) + tQ'_{s-1}(t)\right).$$

We are supposed to conclude from this that $\Delta_1 G_s(\tau_1, \tau_2) = s(s+1)G_s(\tau_1, \tau_2)$ (and indeed the same thing if we replace Δ_1 by the analogous operator Δ_2); but I can't see how this should even be true, since it is not true (as best as I can tell) that $-2((t^2-1)Q''_{s-1}+tQ'_{s-1}) = s(s-1)Q_{s-1}$.

Assuming the claim for the moment, we can think of g_s as a Green's function for the operators $\Delta_m - s(s-1)$ for $m=1,2$, and G_s as an automorphic version given by averaging g_s over Γ. Likewise, we have

$$\Delta_1 E(\tau_1,s) = s(s-1)E(\tau_1,s),$$

and so $G_s(\tau_1,\tau_2)+4\pi E(\tau_1,s)$ is an eigenfunction of Δ_1 with eigenvalue $s(s-1)$, which in the limit as $s\to 1$ is harmonic; and $4\pi(E(\tau_2,s)-\varphi(s))$ is constant in τ_1, and therefore is also harmonic, so the right-hand side is a harmonic function of τ_1. On the other hand for any function h of τ_1 (and possibly τ_2) we have

$$\Delta_1 \log h(\tau_1) = v_1^2\left(\frac{h(\tau_1)h''(\tau_1)-h'(\tau_1)^2}{h(\tau_1)^2}+\frac{h'(\tau_1)^2-h(\tau_1)h''(\tau_1)}{h(\tau_1)^2}\right)=0$$

so the left-hand side is also harmonic; therefore if the two sides agree as $v_1\to\infty$, i.e. they differ by $o(1)$, then they are equal. For the left-hand side, we can use the expansion of the j-invariant

$$j(\tau_1) = e^{-2\pi i\tau_1}+O(1)$$

and so

$$\log|j(\tau_1)-j(\tau_2)|^2 = \log|e^{-2\pi i u_1+2\pi v_1}+O(1)|^2 = 4\pi v_1 + o(1)$$

as $v_1\to\infty$, since we are holding τ_2 constant; and using the Fourier expansions of the terms on the right-hand side we can get the same behavior. $\square$

As in the previous section, we'll assume for simplicity that we're in the generic case where $w_1=w_2=2$, so that

$$\log|J(d_1,d_2)|^2 = \sum_{\tau_1,\tau_2}\log|j(\tau_1)-j(\tau_2)|^2$$

where the sum is taken over classes of τ_m with discriminant d_m for $m=1,2$, or equivalently over all such τ_m in $\Gamma\backslash\mathcal{H}$, where $\mathcal{H}$ is the upper half-plane. In this situation the stabilizers of the τ_m are trivial; therefore

$$\begin{aligned}\sum_{\tau_1,\tau_2} G_s(\tau_1,\tau_2) &= \sum_{\tau_1,\tau_2}\sum_{\gamma\in\Gamma} g_s(\tau_1,\gamma\tau_2)\\ &= \sum_{\tau_1,\tau_2}\sum_{(\gamma_1,\gamma_2)\in\Gamma\backslash(\Gamma\times\Gamma)} g_s(\gamma_1\tau_2,\gamma_2\tau_2)\end{aligned}$$

where the action of Γ on $\Gamma\times\Gamma$ is by $\gamma\cdot(\gamma_1,\gamma_2)=(\gamma\gamma_1,\gamma\gamma_2)$. The set of all $\gamma_m\tau_m$ with τ_m ranging over representatives of $\mathcal{H}$ modulo Γ with discriminant d_m and γ_m ranging over all of Γ is just the set of all $\tau_m\in\mathcal{H}$ with discriminant d_m, so this is

$$\sum_{\substack{(\tau_1,\tau_2)\in\Gamma\backslash\mathcal{H}^2\\ \mathrm{disc}(\tau_m)=d_m}} g_s(\tau_1,\tau_2)$$

with the diagonal action of Γ on $\mathcal{H}^2$. The points τ_m of $\mathcal{H}$ with discriminant d_m correspond to positive-definite quadratic forms of discriminant d_m, with the corresponding form $a_m x^2 + b_m xy + c_m y^2$ such that $a_m\tau_m^2 + b_m\tau_m + c_m = 0$; then we have

$$\tau_m = \frac{-b_m + \sqrt{d_m}}{2a_m}$$

and so $u_m = -\frac{b_m}{2a_m}$ and $v_m = \frac{\sqrt{-d_m}}{2a_m}$. Therefore

$$\frac{(u_1-u_2)^2 + v_1^2 + v_2^2}{2v_1v_2} = \frac{2a_1c_2 + 2a_2c_1 - b_1b_2}{\sqrt{d_1d_2}}$$

since $d_m = b_m^2 - 4a_mc_m$, and so

$$g_s(\tau_1, \tau_2) = -2Q_{s-1}\left(\frac{2a_1c_2 + 2a_2c_1 - b_1b_2}{\sqrt{d_1d_2}}\right).$$

Notice that modulo 2 we have $n := 2a_1c_2 + 2a_2c_1 - b_1b_2 \equiv b_1b_2 \equiv b_1^2b_2^2 \equiv (b_1^2 - 4a_1c_1)(b_2^2 - 4a_2c_2) = d_1d_2$ and $n > \sqrt{d_1d_2}$, so if we define $\rho(n)$ to be the number of pairs positive-definite integral binary quadratic forms $a_mx^2 + b_mxy + c_my^2$ for $m = 1, 2$, modulo the diagonal action of Γ, such that $b_m^2 - 4a_mc_m = d_m$ and $2a_1c_2 + 2a_2c_1 - b_1b_2 = n$ then we have

$$\sum_{\tau_1,\tau_2} G_s(\tau_1, \tau_2) = -2 \sum_{\substack{n > \sqrt{d_1d_2} \\ n \equiv d_1d_2 \pmod 2}} \rho(n) Q_{s-1}\left(\frac{n}{\sqrt{d_1d_2}}\right).$$

For any imaginary quadratic number field K, its Dedekind zeta function decomposes as

$$\zeta_K(s) = \sum_{\mathfrak{n} \subset \mathcal{O}_K} \mathrm{N}(\mathfrak{n})^{-s} = \sum_{\mathfrak{a} \in \mathrm{Cl}(K)} \sum_{\mathfrak{n} \sim \mathfrak{a} \in \mathrm{Cl}(K)} \mathrm{N}(\mathfrak{n})^{-s};$$

call the inner sum $\zeta_{K,\mathfrak{a}}(s)$. For each $\mathfrak{a}$, this correspond to the sum over some lattice $\Lambda_{\mathfrak{a}}$ of discriminant $\operatorname{disc}(K)$ generated by $(1, \tau)$ for some τ

$$\begin{aligned}
\zeta_{K,\mathfrak{a}}(s) &= \sum_{\substack{z \in \Lambda \\ z \neq 0}} \frac{1}{|z|^{2s}} \\
&= \sum_{\substack{c,d \in \mathbb{Z} \\ (c,d) \neq (0,0)}} \frac{1}{|c\tau + d|^{2s}} \\
&= \sum_{k=1}^{\infty} \frac{1}{k^{2s}} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d) = 1}} \frac{1}{|c\tau + d|^{2s}} \\
&= \zeta(2s)\, \mathrm{Im}(\tau)^{-s} E(\tau, s),
\end{aligned}$$

and since τ satisfies a quadratic integral equation of discriminant $\operatorname{disc}(K)$ it has imaginary part $\frac{1}{2}\sqrt{-\operatorname{disc}(K)}$ (given our assumptions). In particular, if $K_m = \mathbb{Q}(\sqrt{d_m})$ is an imaginary quadratic field of discriminant d_m and $\mathfrak{a}$ is the ideal class corresponding to τ_m then

$$\zeta_{K_m,\mathfrak{a}}(s) = \zeta(2s)\left(\frac{d_m}{4}\right)^{-s/2} E(\tau_m, s).$$

Summing over all choices of τ_m up to the action of Γ is the same thing as summing over all the $\mathfrak{a}$, and so

$$\sum_{\tau_m} E(\tau_m, s) = \zeta(2s)^{-1} \left(\frac{d_m}{4}\right)^{s/2} \zeta_{K_m}(s).$$

Since there are h_m choices of τ_m, where h_m is the class number of K_m, we have

$$\sum_{\tau_1,\tau_2} (E(\tau_1, s) + E(\tau_2, s)) = \zeta(2s)^{-1} \left(\left(\frac{d_1}{4}\right)^{s/2} \zeta_{K_1}(s) h_2 + h_1 \left(\frac{d_2}{4}\right)^{s/2} \zeta_{K_2}(s)\right).$$

Therefore we have from Proposition 3.1 the following.

Proposition 3.2. *If K_1 and K_2 are imaginary quadratic number fields of relatively prime discriminants $d_1, d_2 < -4$, then*

$$\begin{aligned}\log |J(d_1, d_2)|^2 = \lim_{s\to 1} \Bigg(&-2 \sum_{\substack{n > \sqrt{d_1 d_2} \\ n \equiv d_1 d_2 \pmod 2}} \rho(n) Q_{s-1}\left(\frac{n}{\sqrt{d_1 d_2}}\right) \\ &+ \frac{4\pi}{\zeta(2s)} \left(\left(\frac{d_1}{4}\right)^{s/2} \zeta_{K_1}(s) h_2 + h_1 \left(\frac{d_2}{4}\right)^{s/2} \zeta_{K_2}(s)\right) \\ &- 4\pi h_1 h_2 \varphi(s)\Bigg) - 24 h_1 h_2.\end{aligned}$$

Note that in this form by making the obvious simplification $\zeta(2s) \to \zeta(2) = \frac{\pi^2}{6}$ and using the class number formula for imaginary quadratic fields, we can see directly that the limit exists as $s \to 1$ (and by being more careful can get a somewhat simpler formula for it, but we will end up approaching this formula from the other direction).

The hardest part of this formula to deal with is the first, since we don't know how to deal with $\rho(n)$; we deal with this using the following lemma.

Lemma 3.3. *Let $\epsilon(n)$ be defined as above. Then*

$$\rho(n) = \sum_{d \mid \frac{n^2 - d_1 d_2}{4}} \epsilon(d).$$

We now turn our attention to the right-hand side of the formula in Theorem 1.1 (or rather its logarithm); we want to massage it into something resembling our formula for the left-hand side. This side appears to be extremely messy, so I'll give only a very rough outline; the upshot is going to be that the expression ends up being miraculously equal to the formula in Proposition 3.2, upon making the substitution from Lemma 3.3.

The idea is as follows: the logarithm of the right-hand side is

$$S := \sum_{\substack{x^2 < d_1 d_2 \\ x^2 \equiv d_1 d_2 \pmod 4}} \sum_{n \mid \frac{d_1 d_2 - x^2}{4}} \epsilon(n) \log n.$$

The idea is that if we were to replace the $\log n$ term with n^s, this would look something like (the digonal of) an Eisenstein series E, twisted by ϵ; to go from n^s to $\log n$, we differentiate with respect to s and evaluate at $s = 0$, so our goal should be to find a suitable Eisenstein series $E_s(\tau_1, \tau_2)$ with $\frac{\partial}{\partial s}\big|_{s=0} E_s(z, z)$ related in some way to S. We can find one whose first Fourier coefficient is equal to S, up to some relatively easy quantity. We can bound the growth of this Eisenstein series logarithmically in terms of certain values of L-functions, and a lemma extending Sturm's techniques on holomorphic projection allows us to relate an integral of the first Fourier coefficient to the constants governing a logarithmic bound on growth. This lets us write S in terms of (a limit of) a certain integral (of the remaining part of the first Fourier coefficient) and various L-function values; the integral, after some massaging, yields the

$$-2 \sum_{n > \sqrt{d_1 d_2} n \equiv d_1 d_2 \pmod 2} \sum_{d \mid \frac{n^2 - d_1 d_2}{4}} \epsilon(d) Q_{s-1}\left(\frac{n}{\sqrt{d_1 d_2}}\right)$$

term, and the remaining terms come from the L-functions after applying the class number formula for imaginary quadratic number fields.

References

[1] David R Dorman. Special values of the elliptic modular function and factorization formulae. *Journal für die reine und angewandte Mathematik*, 1988(383):207–220, 1988.

[2] Benedict H Gross. *Arithmetic on elliptic curves with complex multiplication*, volume 776. Springer, 2006.

[3] Benedict H Gross and Don B Zagier. On singular moduli. *Journal für die reine und angewandte Mathematik*, 355:191–220, 1984.

[4] Jean-Pierre Serre and John Tate. Good reduction of abelian varieties. *Annals of Mathematics*, pages 492–517, 1968.