Heegner points, heights, and L-functions

akrishna168 • 9 Jul 2026

Introduction

This post is a tour of the Gross–Zagier theorem, whose core assertion is that the Néron–Tate height of a Heegner point equals, up to an explicit nonzero constant, the first derivative of an L-function at the center of its functional equation:

$$\hat{h}(P_K) = C(E, K, \varphi)\, L'(E/K, 1).$$

Here, $E/\mathbb{Q}$ is a modular elliptic curve, K an imaginary quadratic field satisfying a Heegner hypothesis, $P_K \in E(K)$ a Heegner point, and $\hat{h}$ the Néron–Tate height. The left side is arithmetic and measures the size of a rational point; the right side is the leading Taylor coefficient of an L-series at its center. Astonishingly, these two real numbers coincide, and the proof shows why: both decompose into local contributions, one per place, which agree place by place. The chain of ideas is

$$\text{CM points on modular curves} \longrightarrow \text{Heegner divisors} \longrightarrow \text{height pairings} \leftarrow$$

Throughout, write $\mathrm{Sha}(E/K)$ for the Tate–Shafarevich group, $w(\cdot)$ for a global root number, χ_D for the quadratic Kronecker character of $K = \mathbb{Q}(\sqrt{D})$, and $d\mu(z) = y^{-2}\,dx\,dy$ for the hyperbolic measure.

---

1. From BSD to Heegner points

1.1 The rank problem

Let $E/\mathbb{Q}$ be an elliptic curve. By the Mordell–Weil theorem, the group of rational points is finitely generated,

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\mathrm{tors}} \oplus \mathbb{Z}^r,$$

and the integer r is the Mordell–Weil rank. The Birch–Swinnerton-Dyer conjecture predicts r analytically through the Hasse–Weil L-function

$$L(E, s) = \prod_p L_p(E, s), \qquad L_p(E, s) = \left(1 - a_p p^{-s} + p^{1-2s}\right)^{-1} \quad (p \nmid N),$$

where $a_p = p + 1 - \#E(\mathbb{F}_p)$ and N is the conductor. The Euler product converges for $\mathrm{Re}(s) > 3/2$, and by modularity continues to an entire function with functional equation $s \leftrightarrow 2 - s$, center $s = 1$. The conjecture asserts

$$\operatorname{rank} E(\mathbb{Q}) = \operatorname{ord}_{s=1} L(E, s),$$

the right side being the analytic rank. Gross–Zagier addresses analytic rank one by producing an explicit point and measuring it.

1.2 Passing to an imaginary quadratic field

It is technically advantageous to work over an imaginary quadratic field $K = \mathbb{Q}(\sqrt{D})$ of discriminant $D < 0$. The base-changed curve E/K has the L-function

$$L(E/K, s) = L(E, s)\, L(E, \chi_K, s),$$

where $\chi_K = \chi_D$ is the quadratic character cutting out K and $L(E, \chi_K, s) = L(E^{(D)}, s)$ is the L-function of the quadratic twist $E^{(D)}$ by D. This factorization is the restriction to G_K of the Galois representation ρ_E of E: $\mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}}(\rho_E|_{G_K}) = \rho_E \oplus (\rho_E \otimes \chi_D)$.

The center is $s = 1$. When the global root number $w(E/K) = -1$, the completed L-function is odd under $s \mapsto 2 - s$, so its central value vanishes and the first candidate invariant is the derivative. A natural hypothesis on K results in this sign.

1.3 The Heegner hypothesis and the sign

An imaginary quadratic field K of discriminant D satisfies the Heegner hypothesis for N if every prime $\ell \mid N$ splits in K. When $(D, N) = 1$, this is equivalent to D being a square modulo $4N$, and it guarantees an ideal $\mathfrak{n} \subset \mathcal{O}_K$ with $\mathcal{O}_K/\mathfrak{n} \cong \mathbb{Z}/N\mathbb{Z}$.

Indeed, each $\ell \mid N$ splits iff D is a nonzero square mod ℓ (with the analogous condition at 2); assembling these by the Chinese remainder theorem gives r with $r^2 \equiv D \pmod{4N}$, and $\mathfrak{n} = (\ell, \frac{r + \sqrt{D}}{2})$ pieced together over $\ell \mid N$ has cyclic quotient.

To determine the sign, we record the local root numbers, using the theory of epsilon factors.

Local root numbers (Deligne, Rohrlich). Let $E/\mathbb{Q}_p$ be an elliptic curve, and let χ be a quadratic character of $\mathbb{Q}_p^\times$. Write $w_p(E) \in \{\pm 1\}$ for the local root number of $E/\mathbb{Q}_p$. Then: 1. if E has good reduction at p and χ is unramified, then $w_p(E) = 1$ and $w_p(E \otimes \chi) = 1$; 2. at the archimedean place,

$w_\infty(E) = -1$, and $w_\infty(E \otimes \chi) = -1$ for every quadratic χ, since the archimedean root number depends only on the Hodge structure, which the twist preserves; 3. if χ is unramified at p (equivalently $p \nmid \operatorname{cond}\chi$) and E has conductor exponent f_p at p, then

$$w_p(E \otimes \chi) = \chi_p(p)^{f_p}\, w_p(E),$$

where $\chi_p(p) = \chi(p) \in \{\pm 1\}$ records the value of the unramified character on a uniformizer.

Statement (3) is the local incarnation of the global twisting formula: it says that twisting by a character unramified at p multiplies the local root number by the value of the character on the local conductor. Assembling the local values gives the global sign.

Sign of the functional equation. Let $E/\mathbb{Q}$ have conductor N, and let K be imaginary quadratic of discriminant D with $(D, N) = 1$ satisfying the Heegner hypothesis for N. Then $w(E/K) = -1$, hence $L(E/K, 1) = 0$.

Proof. By the factorization of Section 1.2, $w(E/K) = w(E)\, w(E^{(D)})$, and since $w(E) = \pm 1$ we have $w(E/K) = w(E)^2 \left(w(E^{(D)})/w(E)\right) = w(E^{(D)})/w(E)$. We compute the ratio place by place using the local root numbers, noting that $E^{(D)} = E \otimes \chi_D$ and that χ_D has conductor $|D|$ coprime to N.

At a finite prime $p \nmid ND$: both E and $E^{(D)}$ have good reduction and χ_D is unramified, so by (1) the local factors are 1 and contribute nothing to the ratio.

At the archimedean place: by (2), $w_\infty(E) = w_\infty(E^{(D)}) = -1$, so the ratio contributes $(-1)/(-1) = +1$.

At a prime $p \mid D$ (so $p \nmid N$, and E has good reduction there while χ_D is ramified): the twist by a ramified quadratic character at a good prime multiplies the local root number by the local root number of χ_D at p, and the product of these local character root numbers over $p \mid D$, together with the archimedean sign of χ_D, assembles into the global root number of χ_D, which for the quadratic character of an imaginary quadratic field equals $\chi_D(-1) = -1$ times a product of split local signs. We will track this below via the aggregate identity.

At a prime $p \mid N$ (so $p \nmid D$, χ_D unramified at p): by (3),

$$\frac{w_p(E^{(D)})}{w_p(E)} = \chi_D(p)^{f_p}.$$

Multiplying the finite ratios and folding in the aggregated ramified and archimedean contributions of χ_D, one obtains the closed form of the twisting formula

$$\frac{w(E^{(D)})}{w(E)} = \chi_D(-N) = \chi_D(-1)\prod_{p\mid N}\chi_D(p)^{v_p(N)}.$$

Now evaluate. Since $D<0$, the character χ_D is odd, so $\chi_D(-1)=-1$. Since every prime $p \mid N$ splits in K, the decomposition law gives $\chi_D(p)=+1$ for each such p, hence $\prod_{p\mid N}\chi_D(p)^{v_p(N)}=+1$. Therefore

$$w(E/K) = \frac{w(E^{(D)})}{w(E)} = \chi_D(-N) = (-1)\cdot(+1) = -1.$$

Consequently, the completed L-function satisfies $\Lambda(E/K,s) = -\Lambda(E/K,2-s)$. Setting $s=1$ gives $\Lambda(E/K,1) = -\Lambda(E/K,1)$, so $\Lambda(E/K,1)=0$, and since the archimedean factor is nonzero at $s=1$, also $L(E/K,1)=0$. □

The vanishing of the central value is what makes the derivative the first informative quantity, and it is this derivative that Gross–Zagier equates to a height.

Gross–Zagier, rank-one form. Let $E/\mathbb{Q}$ be a modular elliptic curve of conductor N with modular parametrization $\varphi : X_0(N) \to E$, and let K be imaginary quadratic satisfying the Heegner hypothesis for N. There is a canonically constructed Heegner point $P_K \in E(K)$ with

$$\hat{h}(P_K) = C(E,K,\varphi)\, L'(E/K,1),$$

where $C(E,K,\varphi)>0$ is an explicit constant built from the Petersson norm of the newform of E, the discriminant $|D|$, and the degree of φ.

The immediate consequence is the equivalence

$$L'(E/K,1)\neq 0 \iff \hat{h}(P_K)\neq 0 \iff P_K \text{ has infinite order},$$

whose rightmost link is the torsion criterion proved in Section 3. In what follows, we describe the construction of P_K and prove the formula.

---

2. Modular curves, CM points, and the construction of Heegner points

2.1 The modular curve as a moduli space

Fix $N \geq 1$. The open modular curve $Y_0(N)$ is the coarse moduli space of pairs (A, C) consisting of an elliptic curve A and a cyclic subgroup $C \subset A$ of order N, and two pairs are identified when related by an isomorphism of elliptic

curves carrying one subgroup to the other. Adjoining finitely many cusps compactifies $Y_0(N)$ to the smooth projective curve $X_0(N)$. Over the complex numbers, there is the uniformization

$$X_0(N)(\mathbb{C}) = \Gamma_0(N)\backslash\mathfrak{H}^*, \qquad \Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : N \mid c \right\},$$

where $\mathfrak{H} = \{\mathrm{Im}(z) > 0\}$ and $\mathfrak{H}^* = \mathfrak{H} \cup \mathbb{P}^1(\mathbb{Q})$ adjoins the cusps. The dictionary sends $\tau \in \mathfrak{H}$ to the lattice $\Lambda_\tau = \mathbb{Z}\tau + \mathbb{Z}$, the curve $A_\tau = \mathbb{C}/\Lambda_\tau$, and the cyclic subgroup $C_\tau = \frac{1}{N}\mathbb{Z}/\mathbb{Z}$; two points give isomorphic pairs iff $\tau' = \gamma\tau$ with $\gamma \in \Gamma_0(N)$.

The curve $X_0(N)$ has a model over $\mathbb{Q}$ whose noncuspidal points classify pairs (A, C). By modularity, $E/\mathbb{Q}$ of conductor N receives a nonconstant $\mathbb{Q}$-morphism

$$\varphi : X_0(N) \to E,$$

normalized with $\varphi(\infty) = O$, with $\varphi^*\omega_E = c_\varphi \cdot 2\pi i\, f(\tau)\, d\tau$ for a nonzero rational c_φ, where $f = \sum_{n \geq 1} a_n q^n$ is the newform of E and $q = e^{2\pi i \tau}$. Albanese functoriality gives a surjection

$$\varphi_* : J_0(N) = \mathrm{Pic}^0(X_0(N)) \longrightarrow E.$$

Rational points on E arise by forming degree-zero divisors from algebraic points on $X_0(N)$ and pushing through φ_*.

2.2 CM points

Let K be imaginary quadratic with ring of integers $\mathcal{O}_K$. A fractional ideal $\mathfrak{a} \subset K$ is a rank-two $\mathbb{Z}$-lattice in $\mathbb{C}$, so $A_\mathfrak{a} = \mathbb{C}/\mathfrak{a}$ is an elliptic curve, and multiplication by $\alpha \in \mathcal{O}_K$ preserves $\mathfrak{a}$, giving $\mathcal{O}_K \hookrightarrow \mathrm{End}(A_\mathfrak{a})$; this is an equality when $\mathfrak{a}$ has multiplier ring $\mathcal{O}_K$, so $A_\mathfrak{a}$ has CM by $\mathcal{O}_K$. The class $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O}_K)$ determines $A_\mathfrak{a}$ up to isomorphism ($\mathbb{C}/\mathfrak{a} \cong \mathbb{C}/\lambda\mathfrak{a}$ via multiplication by λ), so there are $h_K = \#\,\mathrm{Cl}(\mathcal{O}_K)$ curves with CM by $\mathcal{O}_K$.

2.3 Adding level structure: the Heegner points

To lift $A_\mathfrak{a}$ to a point of $X_0(N)$, we need a cyclic subgroup of order N. Fix once and for all an ideal $\mathfrak{n} \subset \mathcal{O}_K$ with $\mathcal{O}_K/\mathfrak{n} \cong \mathbb{Z}/N\mathbb{Z}$, which exists by the Heegner hypothesis. For each fractional ideal $\mathfrak{a}$, set

$$C_\mathfrak{a} = \mathfrak{n}^{-1}\mathfrak{a}/\mathfrak{a} \;\subset\; \mathbb{C}/\mathfrak{a} = A_\mathfrak{a}.$$

Lemma (structure of $C_{\mathfrak a}$). The subgroup $C_\mathfrak{a}$ is cyclic of order N, and it is stable under the CM action of $\mathcal{O}_K$.

Proof. Since $\mathfrak{a} \subseteq \mathfrak{n}^{-1}\mathfrak{a}$ is an inclusion of rank-two lattices, the quotient $\mathfrak{n}^{-1}\mathfrak{a}/\mathfrak{a}$ is finite. Multiplication by $\mathfrak{a}^{-1}$ is an $\mathcal{O}_K$-linear isomorphism

$$\mathfrak{n}^{-1}\mathfrak{a}/\mathfrak{a} \xrightarrow{\sim} \mathfrak{n}^{-1}/\mathcal{O}_K,$$

so we analyze $\mathfrak{n}^{-1}/\mathcal{O}_K$. The multiplication map

$$\langle\ ,\ \rangle : \mathfrak{n}^{-1}/\mathcal{O}_K \times \mathcal{O}_K/\mathfrak{n} \longrightarrow K/\mathcal{O}_K, \qquad (x, y) \mapsto xy \bmod \mathcal{O}_K,$$

is well-defined because $x\mathfrak{n} \subseteq \mathcal{O}_K$ for $x \in \mathfrak{n}^{-1}$, and it is $\mathcal{O}_K$-bilinear. It is a perfect pairing: for fixed nonzero $x \in \mathfrak{n}^{-1}/\mathcal{O}_K$ there is y with $xy \notin \mathcal{O}_K$ because $x \notin \mathcal{O}_K$, and symmetrically, so neither side has a nonzero radical. A perfect pairing of finite abelian groups identifies each factor with the Pontryagin dual $\mathrm{Hom}(-, \mathbb{Q}/\mathbb{Z})$ of the other, and the Pontryagin dual of the cyclic group $\mathcal{O}_K/\mathfrak{n} \cong \mathbb{Z}/N\mathbb{Z}$ is cyclic of order N. Hence $\mathfrak{n}^{-1}/\mathcal{O}_K$, and therefore $C_\mathfrak{a}$, is cyclic of order N.

Stability under $\mathcal{O}_K$ is immediate: for $\alpha \in \mathcal{O}_K$, multiplication by α carries $\mathfrak{n}^{-1}\mathfrak{a}$ into itself and $\mathfrak{a}$ into itself, so it preserves the quotient $C_\mathfrak{a}$. □

The pair $(A_\mathfrak{a}, C_\mathfrak{a})$ defines a point

$$x_\mathfrak{a} \in X_0(N)(\mathbb{C}),$$

the Heegner point of discriminant D attached to $[\mathfrak{a}]$ and $\mathfrak{n}$. In the upper-half-plane model, a basis of $\mathfrak{a}$ adapted to $\mathfrak{n}$ represents $x_\mathfrak{a}$ by $\tau_\mathfrak{a} \in \mathfrak{H}$ satisfying $a\tau^2 + b\tau + c = 0$ with $b^2 - 4ac = D$ and $N \mid a$, the condition $N \mid a$ giving $\Gamma_0(N)$-compatibility. These are the classical Heegner points.

2.4 Fields of definition and the Galois action

The arithmetic of the $x_\mathfrak{a}$ is governed by complex multiplication, whose relevant statement we cite.

Main theorem of complex multiplication. Let H be the Hilbert class field of K, and let $\mathrm{rec} : \mathrm{Cl}(\mathcal{O}_K) \xrightarrow{\sim} \mathrm{Gal}(H/K)$ be the Artin reciprocity isomorphism, $[\mathfrak{b}] \mapsto \sigma_\mathfrak{b}$. Then: 1. every elliptic curve $A_\mathfrak{a}$ with CM by $\mathcal{O}_K$ has a model over H, and $j(A_\mathfrak{a})$ generates H over K; 2. the Galois conjugate $\sigma_\mathfrak{b}(A_\mathfrak{a})$ is isomorphic to $A_{\mathfrak{a}\mathfrak{b}^{-1}}$, compatibly with all torsion, meaning there is an isogeny $A_\mathfrak{a} \to A_{\mathfrak{a}\mathfrak{b}^{-1}}$ inducing $\sigma_\mathfrak{b}$ on torsion points after the class-field-theoretic identification of Frobenius.

From this, we deduce the field of definition and Galois action of the CM points on $X_0(N)$.

Shimura reciprocity for Heegner points. Each Heegner point $x_\mathfrak{a}$ is defined over the Hilbert class field H, and for $\sigma_\mathfrak{b} \in \mathrm{Gal}(H/K)$,

$$\sigma_{\mathfrak{b}}(x_{\mathfrak{a}}) = x_{\mathfrak{a}\mathfrak{b}^{-1}}.$$

In particular $\mathrm{Gal}(H/K)$ acts simply transitively on the set $\{x_{\mathfrak{a}}\}_{[\mathfrak{a}]}$ of h_K Heegner points.

Proof. The point $x_{\mathfrak{a}}$ records the pair $(A_{\mathfrak{a}}, C_{\mathfrak{a}})$. By part (1) of the cited input, $A_{\mathfrak{a}}$ is defined over H. The subgroup $C_{\mathfrak{a}} = \mathfrak{n}^{-1}\mathfrak{a}/\mathfrak{a}$ is cut out by the ideal $\mathfrak{n} \subset \mathcal{O}_K$, which is a fixed rational structure independent of the ideal class; applying part (2), the isogeny realizing $\sigma_{\mathfrak{b}}$ carries $\mathfrak{n}$-torsion to $\mathfrak{n}$-torsion, hence carries $C_{\mathfrak{a}}$ to $C_{\mathfrak{a}\mathfrak{b}^{-1}}$. Therefore,

$$\sigma_{\mathfrak{b}}(A_{\mathfrak{a}}, C_{\mathfrak{a}}) = (A_{\mathfrak{a}\mathfrak{b}^{-1}}, C_{\mathfrak{a}\mathfrak{b}^{-1}}),$$

which is the moduli-theoretic statement $\sigma_{\mathfrak{b}}(x_{\mathfrak{a}}) = x_{\mathfrak{a}\mathfrak{b}^{-1}}$. Because $\{x_{\mathfrak{a}}\}$ is indexed by $\mathrm{Cl}(\mathcal{O}_K)$ and multiplication by $[\mathfrak{b}^{-1}]$ acts simply transitively on the class group, the induced action on Heegner points is simply transitive. In particular each $x_{\mathfrak{a}}$ has exactly h_K Galois conjugates, all Heegner points, so its field of definition is the fixed field of the trivial subgroup, namely H itself when $\mathrm{Cl}(\mathcal{O}_K)$ is nontrivial, and $H = K$ when $h_K = 1$. □

2.5 The Heegner divisor and the Heegner point

Form the divisor

$$Z_K = \sum_{[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O}_K)} (x_{\mathfrak{a}})$$

on $X_0(N)$, of degree h_K. By Shimura reciprocity the Galois group $\mathrm{Gal}(H/K)$ permutes the summands transitively, so Z_K is fixed by $\mathrm{Gal}(H/K)$ and therefore defined over K. The cusp ∞ is defined over $\mathbb{Q}$, so subtracting it produces a degree-zero divisor class over K,

$$y_K = Z_K - h_K(\infty) \in J_0(N)(K).$$

Pushing through the modular parametrization gives the Heegner point on E,

$$P_K = \varphi_*(y_K) = \sum_{[\mathfrak{a}]} \varphi(x_{\mathfrak{a}}) \in E(K),$$

using $\varphi(\infty) = O$. Equivalently, working with a single class $c = (x_1) - (\infty) \in J_0(N)(H)$ for the trivial ideal class and setting $P = \varphi_*(c) \in E(H)$, one has

$$P_K = \mathrm{Tr}_{H/K}(P) = \sum_{\sigma \in \mathrm{Gal}(H/K)} P^{\sigma},$$

since φ_* is Galois-equivariant and the orbit of x_1 is $\{x_{\mathfrak{a}}\}$. The rest of the post computes $\hat{h}(P_K)$.

3. Heights as arithmetic intersection numbers

3.1 The canonical height exists and is quadratic

Let F be a number field and A/F an abelian variety with symmetric ample L; for E take $L = \mathcal{O}(2(O))$, so the naive height is $h(P) = h_{\mathbb{P}^1}(x(P))$. We cite the following consequence of the theorem of the cube.

Quasi-quadraticity of the naive height. For a symmetric ample L on an abelian variety, the naive height $h = h_L$ satisfies, uniformly in P, Q,

$$h(P+Q)+h(P-Q) = 2h(P)+2h(Q)+O(1), \qquad h([m]P) = m^2 h(P)+O(1).$$

Existence and quadraticity of the Néron–Tate height. The limit

$$\hat{h}(P) = \lim_{n\to\infty} \frac{h([2^n]P)}{4^n}$$

exists for every $P \in A(\overline{F})$, differs from h by a bounded function, and defines a quadratic form: the pairing

$$\langle P, Q\rangle = \tfrac{1}{2}\big(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)\big)$$

is bilinear, and $\hat{h}([m]P) = m^2\hat{h}(P)$ for all $m \in \mathbb{Z}$.

Proof. By the cited estimate with $Q = P$, there is a constant C with $|h([2]P) - 4h(P)| \le C$ for all P. Consider the sequence $u_n = 4^{-n}h([2^n]P)$. For $n \ge 0$,

$$|u_{n+1} - u_n| = 4^{-(n+1)}\big|h([2^{n+1}]P) - 4h([2^n]P)\big| \le 4^{-(n+1)}C.$$

Hence, for $m > n$,

$$|u_m - u_n| \le \sum_{k=n}^{m-1} |u_{k+1} - u_k| \le \sum_{k=n}^{m-1} 4^{-(k+1)}C \le \frac{C}{3}\,4^{-n},$$

so (u_n) is Cauchy and the limit $\hat{h}(P)$ exists. Taking $n = 0$ and $m \to \infty$ gives $|\hat{h}(P) - h(P)| \le C/3$, so $\hat{h} - h$ is bounded.

The canonical height inherits a parallelogram law. Applying the cited relation to $[2^n]P$ and $[2^n]Q$, dividing by 4^n, and letting $n \to \infty$, the $O(1)$ term is killed by the factor 4^{-n}, leaving

$$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q).$$

Also, $\hat{h}([2]P) = 4\hat{h}(P)$ directly from the definition, and $\hat{h}(-P) = \hat{h}(P)$ because L is symmetric. A real-valued function on an abelian group satisfying the parallelogram law above and $\hat{h}(0) = 0$ is a quadratic form: define $\langle P, Q\rangle$ by

the displayed polarization; then $(*)$ is equivalent to biadditivity of $\langle\ ,\ \rangle$ in each variable. Symmetry is inbuilt, and $\langle P, P\rangle = \hat{h}(P)$. Biadditivity gives $\hat{h}([m]P) = \langle [m]P, [m]P\rangle = m^2\langle P, P\rangle = m^2\hat{h}(P)$ for all integers m. □

3.2 Positivity and the torsion criterion

Positivity of the height, and the certification of infinite order, rest on Northcott finiteness.

Northcott finiteness. For every real B and integer d, the set of algebraic numbers α with $[\mathbb{Q}(\alpha) : \mathbb{Q}] \le d$ and (absolute logarithmic Weil) height $h(\alpha) \le B$ is finite. Consequently, for a fixed number field F, the set $\{P \in E(F) : h(P) \le B\}$ is finite.

Proof. Let α be algebraic of degree $e \le d$ with conjugates $\alpha_1, \dots, \alpha_e$ and minimal polynomial $m_\alpha(T) = \prod_i (T - \alpha_i) = \sum_k (-1)^k s_k T^{e-k}$, where s_k is the k-th elementary symmetric function of the conjugates. The Mahler measure $M(\alpha) = \exp(e \cdot h(\alpha))$ dominates each symmetric function: by the triangle inequality applied to the coefficients, $|s_k| \le \binom{e}{k} M(\alpha) \le 2^e M(\alpha) = 2^e e^{eh(\alpha)} \le 2^d e^{dB}$. Thus, each coefficient of m_α is a rational integer (when α is an algebraic integer; in general one clears a bounded denominator similarly bounded by the height) of absolute value at most a fixed bound depending only on d, B. There are finitely many integer polynomials of degree $\le d$ with coefficients so bounded, hence finitely many possible m_α, hence finitely many α.

For the elliptic curve statement, each $P \in E(F)$ has $x(P) \in F$ of degree $\le [F : \mathbb{Q}]$ and height $h(x(P)) = h(P) \le B$; by the first part there are finitely many possible x-coordinates, and each corresponds to at most two points P. Hence, the set is finite. □

Torsion criterion. For $P \in E(F)$, one has $\hat{h}(P) = 0$ if and only if P is torsion. Equivalently, $\hat{h}$ is positive definite on $E(F) \otimes \mathbb{R}$.

Proof. If P is torsion, $[m]P = O$ for some $m \ge 1$, so $m^2\hat{h}(P) = \hat{h}([m]P) = 0$ and $\hat{h}(P) = 0$.

Conversely, suppose $\hat{h}(P) = 0$. Then $\hat{h}([n]P) = n^2\hat{h}(P) = 0$ for all n. Since $|h - \hat{h}| \le C'$ is bounded, $h([n]P) \le C'$ for all n, so the infinite set $\{[n]P : n \in \mathbb{Z}\}$ lies in the finite set $\{Q \in E(F) : h(Q) \le C'\}$ of Northcott's theorem. Therefore two distinct multiples coincide, $[n_1]P = [n_2]P$ with $n_1 \ne n_2$, giving $[n_1 - n_2]P = O$, so P is torsion. Positive-definiteness on $E(F) \otimes \mathbb{R}$ follows because $\hat{h}$ is a quadratic form vanishing exactly on the torsion subgroup, which becomes zero in $E(F) \otimes \mathbb{R}$. □

3.3 Local decomposition: Néron's local heights

The global height pairing on the Jacobian decomposes into local terms. We state Néron's theorem, take its existence for granted, and prove the well-definedness we need through Weil reciprocity and the product formula.

For a smooth projective geometrically connected curve X/F with Jacobian $J = \mathrm{Pic}^0(X)$, and for each place v of F with normalized absolute value $|\cdot|_v$, write F_v for the completion and $\log q_v$ for the local normalization (q_v the residue cardinality at finite v, and the usual archimedean normalization at $v \mid \infty$).

Néron local heights. There is a unique family of local symbols $\langle\ ,\ \rangle_v$, defined on pairs of degree-zero divisors on X_{F_v} with disjoint support, that is: 1. biadditive and symmetric; 2. continuous in each variable for the v-adic topology; 3. normalized by Weil reciprocity: $\langle D, \mathrm{div}(g)\rangle_v = \log |g(D)|_v$ whenever $\mathrm{supp}\, D \cap \mathrm{supp}\, \mathrm{div}(g) = \varnothing$, where $g(D) = \prod_i g(P_i)^{n_i}$ for $D = \sum_i n_i (P_i)$.

Global height as a sum of local heights. For $a, b \in J(F)$ represented by degree-zero divisors D_a, D_b with disjoint support, the Néron–Tate pairing is

$$\langle a, b\rangle = -\sum_v \langle D_a, D_b\rangle_v,$$

the sum over all places of F, and the right side is independent of the chosen representatives.

Proof of well-definedness. We prove independence of the choice of representatives, which is the crux; identifying the resulting biadditive symmetric function with the Néron–Tate pairing is then the cited characterization of the latter as the unique quadratic form inducing the canonical class.

Suppose D_a and D_a' both represent a, so $D_a' = D_a + \mathrm{div}(g)$ for some rational function $g \in F(X)^\times$. Fix D_b with support disjoint from both. Then by biadditivity and the Weil reciprocity property,

$$\sum_v \langle D_a', D_b\rangle_v - \sum_v \langle D_a, D_b\rangle_v = \sum_v \langle \mathrm{div}(g), D_b\rangle_v = \sum_v \langle D_b, \mathrm{div}(g)\rangle_v = \sum_v \log |$$

Now $D_b = \sum_j m_j (Q_j)$ has degree zero, so $g(D_b) = \prod_j g(Q_j)^{m_j}$ is a well-defined element of $F^\times$ (the degree-zero condition removes the ambiguity of scaling g). By the product formula for the number field F,

$$\sum_v \log |g(D_b)|_v = 0.$$

Therefore, the total sum is unchanged when D_a is replaced by D_a', and by symmetry also when D_b is varied. Hence, $-\sum_v \langle D_a, D_b\rangle_v$ depends only on the classes a, b, is biadditive and symmetric, and is continuous. Therefore, it agrees with the Néron–Tate pairing after normalizing the sign so that the archimedean contributions make the induced quadratic form positive definite, which is the cited final identification. □

3.4 Finite local heights through intersection theory

At a finite place the local symbol is an intersection number on a regular model. Let v be finite with valuation ring $\mathcal{O}_v$, uniformizer π_v, residue field of cardinality q_v, and valuation ord_v, and choose a regular proper flat model $\mathcal{X}/\mathcal{O}_v$ of X_{F_v}. A point $P \in X(F_v)$ extends by properness to a section $\overline{P} : \operatorname{Spec} \mathcal{O}_v \to \mathcal{X}$, a horizontal prime divisor meeting the special fiber in one closed point.

Local intersection multiplicity as proximity. Suppose two sections $\overline{P}, \overline{Q}$ meet at a closed point x of the special fiber, and choose a local coordinate so that $\widehat{\mathcal{O}}_{\mathcal{X},x} \cong \mathcal{O}_v[[u]]$ with $\overline{P} = \{u = 0\}$ and $\overline{Q} = \{u = a\}$ for some $a \in \mathfrak{m}_v = (\pi_v)$. Then, the intersection multiplicity is

$$(\overline{P} \cdot \overline{Q})_x = \operatorname{length}_{\mathcal{O}_v}\big(\mathcal{O}_{\mathcal{X},x}/(I_P + I_Q)\big) = \mathrm{ord}_v(a).$$

Proof. The ideals are $I_P = (u)$ and $I_Q = (u - a)$, so $I_P + I_Q = (u, u - a) = (u, a)$. The quotient is

$$\mathcal{O}_v[[u]]/(u, a) \cong \mathcal{O}_v/(a) \cong \mathcal{O}_v/(\pi_v^{\mathrm{ord}_v(a)}),$$

whose length as an $\mathcal{O}_v$-module is $\mathrm{ord}_v(a)$. Geometrically, the two sections agree to v-adic order exactly $\mathrm{ord}_v(a)$, so intersection multiplicity measures the order of v-adic contact. □

For general degree-zero divisors, one extends by biadditivity and repairs model-independence by a vertical correction, a morsel of linear algebra we carry out.

Existence and uniqueness of the vertical correction. Let the special fiber of $\mathcal{X}$ be $\sum_i m_i \Gamma_i$ with irreducible components Γ_i and multiplicities m_i. Let $\overline{D}$ be the closure of a degree-zero divisor on the generic fiber. Then there is a vertical $\mathbb{Q}$-divisor $V = \sum_i c_i \Gamma_i$, unique up to adding a rational multiple of the whole fiber, such that $(\overline{D} + V) \cdot \Gamma_j = 0$ for every j. The corrected local symbol

$$\langle D, D'\rangle_v = \big((\overline{D} + V) \cdot \overline{D'}\big) \log q_v$$

is independent of the model and satisfies the Néron axioms.

Proof. Let $M = (\Gamma_i \cdot \Gamma_j)_{i,j}$ be the intersection matrix of the special fiber. Two structural facts about M come from the geometry of a proper regular arithmetic surface: first, the fiber class $\mathcal{F} = \sum_i m_i \Gamma_i$ satisfies $\mathcal{F} \cdot \Gamma_j = 0$ for all j, because $\mathcal{F}$ is a fiber of a proper flat map and is algebraically equivalent to a multiple of any other fiber, so it meets each vertical curve in degree zero; hence $Mm = 0$ where $m = (m_i)$, and M is singular with m in its kernel. Second, by the theorem of Zariski (the arithmetic Hodge index theorem for the special fiber), M is negative semidefinite with kernel exactly the line $\mathbb{Q} \cdot m$.

We seek $c = (c_i)$ with $Mc = -r$, where $r = (r_j)$, $r_j = \overline{D} \cdot \Gamma_j$. A vector is in the image of the symmetric matrix M if and only if it is orthogonal to $\ker M = \mathbb{Q} \cdot m$. Compute

$$r \cdot m = \sum_j m_j (\overline{D} \cdot \Gamma_j) = \overline{D} \cdot \Big(\sum_j m_j \Gamma_j \Big) = \overline{D} \cdot \mathcal{F} = \deg(D) = 0,$$

using that $\overline{D} \cdot \mathcal{F}$ equals the degree of D on the generic fiber (a horizontal divisor meets a fiber in its generic degree) and $\deg D = 0$. Thus $-r \perp m$, so $-r \in \operatorname{im} M$ and a solution c exists; it is unique modulo $\ker M = \mathbb{Q} \cdot m$, that is, modulo adding a multiple of the fiber.

Adding a multiple $t\mathcal{F}$ of the fiber to V changes $(\overline{D} + V) \cdot \overline{D'}$ by $t(\mathcal{F} \cdot \overline{D'}) = t \deg(D') = 0$, so the corrected symbol is independent of that ambiguity, hence model-independent after the standard comparison of models by blow-ups. Biadditivity and symmetry are inherited from the intersection product, and the Weil reciprocity normalization is checked by taking $D' = \operatorname{div}(g)$ and using that principal divisors have trivial corrected self-intersection with horizontal classes, which recovers $\log |g(D)|_v$. Thus the corrected symbol satisfies the Néron axioms and agrees with the local height. □

3.5 Archimedean local heights through Green functions

At an archimedean place the local symbol is computed by a Green function on the Riemann surface $X(\mathbb{C})$. Fix a smooth positive $(1,1)$-form μ with $\int_{X(\mathbb{C})} \mu = 1$.

A Green function for a point $Q \in X(\mathbb{C})$ is a smooth function g_Q on $X(\mathbb{C}) \setminus \{Q\}$ with the properties

$$\frac{1}{\pi i}\, \partial\overline{\partial}\, g_Q = \delta_Q - \mu \quad \text{(as currents)}, \qquad \int_{X(\mathbb{C})} g_Q \, \mu = 0,$$

so that near Q in a local coordinate z one has $g_Q(z) = \log |z - z(Q)| + O(1)$.

Archimedean local symbol. For degree-zero divisors $D = \sum_i n_i (P_i)$ and $D' = \sum_j m_j (Q_j)$ with disjoint support, the archimedean local symbol is

$$\langle D, D' \rangle_v = \sum_{i,j} n_i m_j \, g_{Q_j}(P_i),$$

and this is symmetric in D, D' and independent of μ.

Proof. Set $g_{D'} = \sum_j m_j g_{Q_j}$. Because $\deg D' = \sum_j m_j = 0$, the defining equation gives $\frac{1}{\pi i}\partial\overline{\partial} g_{D'} = \delta_{D'} - (\deg D')\mu = \delta_{D'}$, so $g_{D'}$ is a genuine potential for D' with no μ-dependence, and the normalization integral drops out. Evaluating at D gives $\langle D, D' \rangle_v = \sum_i n_i g_{D'}(P_i) = \sum_{i,j} n_i m_j g_{Q_j}(P_i)$. Symmetry is Green's reciprocity: for degree-zero D, D' with disjoint support,

$$\sum_i n_i g_{D'}(P_i) = \sum_j m_j g_D(Q_j),$$

which follows from Stokes' theorem applied to $g_D \, \partial\overline{\partial} g_{D'} - g_{D'} \, \partial\overline{\partial} g_D$, whose integral vanishes because the two currents δ_D and $\delta_{D'}$ are integrated against the complementary potentials symmetrically. Independence of μ is the same degree-zero cancellation. This is the local height at the archimedean place. □

On modular curves, this Green function is the automorphic one, built from the hyperbolic resolvent. Its Legendre-function form is derived in Section 5.

---

4. Local heights at finite places: the quaternionic computation

Gross and Zagier compute the whole family $\langle c, T_m c \rangle$, matched coefficient by coefficient against an analytic kernel. We reduce each finite local height to a count of quaternionic data, down to the deformation-theoretic and quaternionic inputs of the status note.

4.1 Hecke operators as correspondences

The Hecke operator T_m acts on divisors of $X_0(N)$ by summing over cyclic m-isogenies. On moduli,

$$T_m(A, C) = \sum_{\substack{\phi : A \to A' \\ \deg \phi = m,\ \ker \phi \text{ cyclic},\ \ker \phi \cap C = 0}} (A', \phi(C)),$$

the sum over the finitely many cyclic m-isogenies from A whose kernel meets C trivially; equivalently T_m is the correspondence $X_0(N, m)$ with its two projections to $X_0(N)$. For a Heegner point, when $\mathfrak{m} \subset \mathcal{O}_K$ has norm m the isogeny $A_{\mathfrak{a}} \to A_{\mathfrak{m}^{-1}\mathfrak{a}}$ realizes such an isogeny, so Hecke translates of CM points are again CM points, which is the source of the arithmetic below.

4.2 Reduction of CM points and the collision condition

Fix a rational prime ℓ and a place λ of H above ℓ, and extend $X_0(N)$ to a regular model over the ring of integers, so Heegner points reduce to the special fiber over $\overline{\mathbb{F}}_\ell$. Two components $x_{\mathfrak{a}}^{\sigma}$ and $(T_m x_{\mathfrak{b}})^{\tau}$ contribute to the finite local height at λ exactly when their reductions coincide, the contribution being the intersection multiplicity of Section 3.4.

The reduction of a CM elliptic curve $A_{\mathfrak{a}}$ modulo λ depends on the splitting of ℓ in K:

- if ℓ splits in K, the reduction is ordinary, with $\mathrm{End}(\overline{A}_{\mathfrak{a}})$ an order in K containing $\mathcal{O}_K$; distinct Heegner points then reduce to distinct ordinary points unless related by a genuine ℓ-power isogeny, and the collisions are controlled by the ordinary deformation theory;
- if ℓ is inert or ramified in K, the reduction can be supersingular, and $\mathrm{End}(\overline{A}_{\mathfrak{a}})$ is a maximal order in the definite quaternion algebra $B = B_{\ell,\infty}$ ramified exactly at ℓ and ∞.

The supersingular case carries the essential arithmetic: there the endomorphism ring is noncommutative, and $\mathcal{O}_K$ embeds into it in finitely many ways, each recording a way for CM points to collide.

4.3 Intersection multiplicity as a lifting length

The intersection multiplicity of two sections agreeing modulo λ is the length of their largest common infinitesimal deformation, computed by the following input.

Serre–Tate deformation theory. Let $\overline{A}$ be an elliptic curve over a perfect field of characteristic ℓ, and let R be a complete local Artinian ring with residue field that field. The deformations of $\overline{A}$ over R are governed by the deformation of its ℓ-divisible group; in the ordinary case the deformation functor is prorepresentable by a formal torus (the Serre–Tate canonical coordinate q), and a homomorphism $\overline{\beta} : \overline{A}_1 \to \overline{A}_2$ of the reductions lifts to level λ^n if and only if it is compatible with the canonical coordinates to that order. In the supersingular case the analogous statement holds with the endomorphism action of the quaternion order, and the lifting of a given $\overline{\beta}$ to successive infinitesimal orders is obstructed by the quasi-canonical lifting filtration, so that $\overline{\beta}$ lifts modulo λ^n precisely for n up to an explicit level determined by the discriminant of the CM order relative to $\overline{\beta}$.

Local intersection as a homomorphism count. Let A_1, A_2 be CM elliptic curves with the same reduction $\overline{A}$ modulo λ. Then the local intersection multiplicity of the corresponding Heegner sections, restricted to the degree-m Hecke component, is

$$(\overline{x}_1 \cdot \overline{x}_2)_\lambda = \tfrac{1}{2} \sum_{n \geq 1} \#\big\{\overline{\beta} \in \mathrm{Hom}(\overline{A}_1, \overline{A}_2) : \deg \overline{\beta} = m,\ \overline{\beta} \text{ lifts modulo } \lambda^n,\ \overline{\beta} \text{ respe}$$

Proof. By Section 3.4 the intersection multiplicity at λ equals ord_λ of the parameter measuring how far the two sections stay equal, which by definition is the largest n such that the two deformations agree modulo λ^n. A degree-m isogeny between the generic CM curves reduces to a homomorphism $\overline{\beta}$ of degree m between the reductions; conversely, by Serre–Tate, such a $\overline{\beta}$ lifts to an isogeny modulo λ^n exactly when it is compatible with the deformation coordinates to order n. Counting, for each order n, the homomorphisms that persist to that order and summing over n therefore reproduces the total order of contact, since a homomorphism contributing contact exactly to order n_0 is counted once for each $n \leq n_0$, giving total weight n_0, which is the intersection multiplicity it produces. The factor $\frac{1}{2}$ removes the double count from the involution $\overline{\beta} \leftrightarrow -\overline{\beta}$, which give the same unoriented contact. The level-structure compatibility restricts to those $\overline{\beta}$ carrying C_1 to C_2, matching the Hecke condition. □

4.4 Passing to the quaternion algebra: the Deuring correspondence

At a supersingular prime the homomorphism modules live in a quaternion algebra, computed by the Deuring correspondence and Eichler's embedding theory.

Deuring correspondence (cited input). Fix a prime ℓ and let $B = B_{\ell,\infty}$ be the definite quaternion algebra over $\mathbb{Q}$ ramified exactly at ℓ and ∞, with a maximal order $\mathcal{O}_B$; for level N prime to ℓ use an Eichler order $\mathcal{O}_B(N)$ of level N. Then: 1. the supersingular elliptic curves over $\overline{\mathbb{F}}_\ell$ with $\Gamma_0(N)$-level structure are in bijection with the right ideal classes of $\mathcal{O}_B(N)$; 2. for two such curves corresponding to ideal classes with connecting ideal I, the module $\mathrm{Hom}(\overline{A}_1, \overline{A}_2)$ of homomorphisms is identified with I, and the degree of a homomorphism corresponds to the reduced norm nrd divided by $\mathrm{nrd}(I)$; 3. the CM by $\mathcal{O}_K$ on a supersingular curve corresponds to an embedding $\mathcal{O}_K \hookrightarrow \mathcal{O}_B(N)$, optimal in the sense of Section 4.5.

Eichler optimal embedding numbers (cited input, explicit form). Let $\left(\frac{D}{p}\right) \in \{+1, -1, 0\}$ be the Eichler symbol of K at p (equal to $+1$ if p splits in K, -1 if p is inert, 0 if p ramifies). The mass of optimal embeddings of $\mathcal{O}_K$ into an Eichler order of level N in $B_{\ell,\infty}$, meaning the sum over the ideal classes $\mathcal{O}_i$ of the number $m(\mathcal{O}_K, \mathcal{O}_i)$ of optimal embeddings up to unit conjugacy weighted by $1/w_i$ with $w_i = \frac{1}{2}\#\mathcal{O}_i^\times$, factors as a product of local embedding numbers:

$$\sum_i \frac{m(\mathcal{O}_K, \mathcal{O}_i)}{w_i} = \frac{h(D)}{w_K} \left(1 - \left(\tfrac{D}{\ell}\right)\right) \prod_{p \mid N} \left(1 + \left(\tfrac{D}{p}\right)\right),$$

where $h(D) = \#\operatorname{Cl}(\mathcal{O}_K)$ and $w_K = \frac{1}{2}\#\mathcal{O}_K^\times$. The local factor at ℓ is $1 - \left(\frac{D}{\ell}\right)$, which is 2 when ℓ is inert, 1 when ℓ ramifies, and 0 when ℓ splits; thus supersingular CM points, and hence finite local heights, occur exactly at primes ℓ nonsplit in K. Under the Heegner hypothesis each factor at $p \mid N$ is $1 + \left(\frac{D}{p}\right) = 2$.

The lifting length attached to each embedding is governed by the theory of quasi-canonical liftings, which we also cite in explicit form.

Quasi-canonical lifting length (Gross, cited input). Let $\overline{A}$ be a supersingular curve over $\overline{\mathbb{F}}_\ell$ with $\operatorname{End}(\overline{A}) = \mathcal{O}_B$, and let $\iota : \mathcal{O}_K \hookrightarrow \mathcal{O}_B$ be an optimal embedding. For each $n \geq 0$ there is a quasi-canonical lift A_n of level n, an elliptic curve in characteristic zero with CM by the order $\mathcal{O}_{K,n} = \mathbb{Z} + \ell^n \mathcal{O}_K$ of conductor ℓ^n, defined over the ring class field H_n of that conductor, with

$$[H_n : H_0] = \ell^{n-1}\left(\ell - \left(\tfrac{D}{\ell}\right)\right) \quad (n \geq 1), \qquad [H_0 : K] = h(D).$$

Two CM points whose relative position is measured by an endomorphism of conductor exactly ℓ^n have order of contact (intersection multiplicity) modulo λ equal to the level n at which their quasi-canonical lifts first diverge, and the ℓ-adic valuation of the difference of their Serre–Tate parameters is exactly n.

4.5 Optimal embeddings and the closed formula

Let $\mathcal{O}$ be an order in a quaternion algebra B over $\mathbb{Q}$ and let S be an order in an imaginary quadratic field $L \subset B$. An embedding $\iota : L \hookrightarrow B$ with $\iota(L) \cap \mathcal{O} = \iota(S)$ is an optimal embedding of S into $\mathcal{O}$; that is, S is exactly the elements of L that land in $\mathcal{O}$.

To turn the count into an explicit arithmetic function, fix the two optimal embeddings $\iota_1, \iota_2 : \mathcal{O}_K \hookrightarrow \mathcal{O}_B(N)$ coming from the two colliding CM points, and let I be the connecting ideal, so that $\operatorname{Hom}(\overline{A}_1, \overline{A}_2) \cong I$ with $\deg \beta = \operatorname{nrd}(\beta)/\operatorname{nrd}(I)$. The homomorphisms compatible with the CM

structure are those in the orthogonal complement of $\iota_1(\mathcal{O}_K)$ under the reduced-trace pairing; on that rank-two complement (a rank-two lattice, since I has rank four and the CM condition imposes two linear conditions) the reduced norm restricts to a positive-definite binary quadratic form Q_I of discriminant $D \cdot \ell N^2 / \operatorname{nrd}(I)^2$ up to units. Write

$$r_{Q_I}(t) = \#\{\beta \in I^\perp : \operatorname{nrd}(\beta) = t\}$$

for its representation numbers.

Finite local height at a nonsplit prime, explicit form. For ℓ nonsplit in K and $(m, N\ell) = 1$, the finite local height contribution to $\langle c, T_m c\rangle$ at ℓ is

$$\sum_{\lambda|\ell} (\overline{c} \cdot \overline{T_m c})_\lambda \, \log q_\lambda = \Big(\sum_{n\geq 1} r_{Q_I}\big(m \cdot \operatorname{nrd}(I)\big)_{(n)}\Big) \log \ell,$$

where $r_{Q_I}(\cdot)_{(n)}$ counts those β whose associated CM order has conductor exactly ℓ^n, so that each such β carries the quasi-canonical lifting weight n. Equivalently, grouping by the conductor of the order generated by β,

$$\sum_{\lambda|\ell} (\overline{c} \cdot \overline{T_m c})_\lambda \, \log q_\lambda = \Big(\sum_{n\geq 1} n \cdot \big(\rho_n(m) - \rho_{n+1}(m)\big)\Big) \log \ell = \Big(\sum_{n\geq 1} \rho_n(m)\Big) \log \ell,$$

where $\rho_n(m)$ is the number of optimal embeddings of the conductor-ℓ^n order $\mathcal{O}_{K,n}$ compatible with a degree-m homomorphism, computed by the explicit Eichler numbers with D replaced by $\ell^{2n} D$.

Proof. By Section 4.3, the local multiplicity of two colliding CM points is $\sum_{n\geq 1} \#\{\overline{\beta} : \deg \overline{\beta} = m,\ \overline{\beta} \text{ lifts mod } \lambda^n\}$, up to the factor $\frac{1}{2}$ absorbed into the embedding-number normalization. By the Deuring correspondence a degree-m homomorphism is an element $\beta \in I$ with $\operatorname{nrd}(\beta) = m \operatorname{nrd}(I)$, and compatibility with the two CM structures places β in the complement $I^\perp$ on which the reduced norm is the binary form Q_I; so the homomorphisms of degree m are counted by $r_{Q_I}(m \operatorname{nrd}(I))$. By the quasi-canonical lifting input, a given β lifts modulo λ^n precisely when its associated CM order $\mathbb{Z}[\beta] \cap \iota(K)$ has conductor divisible by ℓ^n; the largest such n, the order of contact this β produces, is the ℓ-adic conductor exponent of $\mathbb{Z}[\beta]$. Therefore β contributes weight equal to that exponent, and summing the indicator "lifts mod λ^n" over $n \geq 1$ recovers the exponent as the total weight. Writing $\rho_n(m)$ for the number of degree-m compatible β whose conductor exponent is at least n, the sum of weights is $\sum_{n\geq 1} \rho_n(m)$, which telescopes as displayed. Each $\rho_n(m)$ is itself an Eichler optimal embedding number for the order $\mathcal{O}_{K,n}$ of discriminant $\ell^{2n} D$, hence given explicitly by the product formula with the symbol $\left(\frac{\ell^{2n} D}{\cdot}\right)$; this makes

every term computable. Finally each supersingular closed point λ has residue cardinality a power of ℓ, so $\log q_\lambda \in \mathbb{Z}_{\geq 1} \log \ell$, and the aggregate is the stated multiple of $\log \ell$. $\square$

4.6 The finite side as a generating series

Assembling over m, the finite local heights are the Fourier coefficients of a generating series

$$\Xi_{\text{fin}}(z) = \sum_{m \geq 1} \Big(\sum_{v \nmid \infty} \langle c, T_m c \rangle_v \Big) q^m.$$

By the explicit local height formula, the coefficient at m is $\big(\sum_{n \geq 1} \rho_n(m)\big) \log \ell$ summed over nonsplit ℓ, where $\rho_n(m)$ counts the degree-m homomorphisms intertwining the two CM structures whose associated order has conductor at least ℓ^n. These intertwiners form a rank-two lattice L_n (the space intertwining two embeddings of K into the quaternion algebra is a rank-two $\mathcal{O}_K$-module, since the centralizer of a maximal subfield is the field itself), and the reduced norm restricts to a positive-definite binary quadratic form Q_n on L_n, so

$$\rho_n(m) = \#\{v \in L_n : Q_n(v) = m\} = r_{Q_n}(m), \qquad \vartheta_{Q_n}(z) = \sum_{v \in L_n} q^{Q_n(v)} = \sum_{t \geq 0} r$$

is a weight-one theta series attached to the imaginary quadratic order, in fact a linear combination of the very theta series θ_A of Section 6.1 for the conductor-ℓ^n order. The finite generating series is therefore the $\mathbb{Z}$-linear combination

$$\Xi_{\text{fin}}(z) = \sum_{\ell \text{ nonsplit}} (\log \ell) \sum_{n \geq 1} \vartheta^{+}_{Q_n}(z)$$

of the nonconstant parts of these weight-one theta series, and the sum over lifting levels n is exactly the coefficientwise s-derivative that arises when the weight-one Eisenstein series $E_1(z, s)$ is differentiated at the center, since ∂_s of the Euler factor at ℓ contributes the grading by n with weight $\log \ell$. On the analytic side the nonarchimedean Fourier coefficient of $\partial_s E_1|_{s_0}$ at index m is the s-derivative of a product of local factors $\prod_p W_p(m; s)$, each W_p a local Whittaker function whose value is the local density of representations of m by the reduced-norm form; the logarithmic derivative $\partial_s W_\ell|_{s_0}$ produces the same weight $\log \ell$ and the same grading by n as $\rho_n(m)$. Matching the two coefficientwise is thus an identity of local representation densities against local Whittaker derivatives, and it is the nonarchimedean half of the comparison. The archimedean half is the content of the next section.

5. Local heights at the archimedean place: the hyperbolic Green function

The archimedean local height of Heegner divisors is a sum of values of the Green function on $X_0(N)(\mathbb{C})$, which we derive as the resolvent kernel of the hyperbolic Laplacian, identify with a Legendre function, and match to the analytic kernel.

5.1 The hyperbolic Laplacian on radial functions

On $\mathfrak{H}$ with coordinates $z = x + iy$, the hyperbolic Laplacian is

$$\Delta = -y^2\big(\partial_x^2 + \partial_y^2\big),$$

invariant under $\mathrm{SL}_2(\mathbb{R})$. The geodesic distance $d(z,w)$ between two points is encoded by the point-pair invariant

$$u(z,w) = \frac{|z-w|^2}{y\,y'}, \qquad z = x+iy,\ \ w = x'+iy',$$

through $\cosh d(z,w) = 1 + \frac{1}{2}u(z,w)$. Write $t = 1 + \frac{1}{2}u = \cosh d \in [1,\infty)$. On a radial function $F = F(t)$ the Laplacian acts as an ordinary differential operator.

Radial reduction of the Laplacian. For a radial function $F = F(t)$ with $t = \cosh d(z,w)$,

$$\Delta_z F = -\big(t^2-1\big)F''(t) - 2t\,F'(t).$$

Proof. By $\mathrm{SL}_2(\mathbb{R})$-invariance of both Δ and t, it suffices to compute at a convenient configuration and with w fixed. Place $w = i$ and use geodesic polar coordinates (d,θ) centered at w, in which the hyperbolic metric is $ds^2 = dd^2 + \sinh^2 d\,d\theta^2$ and the Laplacian is

$$\Delta = -\Big(\partial_d^2 + \coth d\,\partial_d + \frac{1}{\sinh^2 d}\partial_\theta^2\Big).$$

A radial function is independent of θ, so

$$\Delta F = -\big(\partial_d^2 + \coth d\,\partial_d\big)F.$$

Now change variables to $t = \cosh d$, with $\frac{dt}{dd} = \sinh d$ and $\sinh^2 d = t^2 - 1$. Then $\partial_d F = \sinh d\,F'(t)$ and

$$\partial_d^2 F = \cosh d\,F'(t) + \sinh^2 d\,F''(t) = t\,F'(t) + (t^2-1)F''(t).$$

Also $\coth d\,\partial_d F = \frac{\cosh d}{\sinh d}\sinh d\,F'(t) = t\,F'(t)$. Adding,

$$\Delta F = -\big[(t^2-1)F''(t) + tF'(t) + tF'(t)\big] = -(t^2-1)F''(t) - 2tF'(t).$$

□

5.2 The free-space Green function and the Legendre equation

We want the resolvent kernel g_s solving $(\Delta - s(s-1))g_s = 0$ off the diagonal, with a log singularity on it and decay at infinity. By the radial reduction, this is an ODE.

Legendre form of the free Green function. A radial solution $F(t)$ of $(\Delta - s(s-1))F = 0$ solves Legendre's differential equation

$$(t^2-1)F''(t) + 2t\,F'(t) - s(s-1)F(t) = 0,$$

whose solution decaying as $t \to \infty$ is the Legendre function of the second kind $Q_{s-1}(t)$. Consequently

$$g_s(z,w) = -2\,Q_{s-1}\big(1+\tfrac{1}{2}u(z,w)\big), \qquad Q_{s-1}(t) = \int_0^\infty \big(t+\sqrt{t^2-1}\,\cosh\xi\big)^{-s}\,d\xi$$

is a free-space Green function for $\Delta - s(s-1)$, holomorphic in s for $\mathrm{Re}(s) > 0$, $t > 1$.

Proof. Substituting the radial reduction into $(\Delta - s(s-1))F = 0$ gives

$$-(t^2-1)F'' - 2tF' - s(s-1)F = 0,$$

that is, $(t^2-1)F'' + 2tF' + s(s-1)F = 0$. Writing $\nu = s-1$ so that $s(s-1) = \nu(\nu+1)$ turns this into the standard Legendre equation

$$(t^2-1)F'' + 2tF' - \nu(\nu+1)F = 0$$

of degree $\nu = s-1$ (the sign of the zeroth-order term matches after $s(s-1) = \nu(\nu+1)$). Legendre's equation has a two-dimensional solution space spanned by $P_\nu(t)$ and $Q_\nu(t)$; the first grows and the second decays as $t \to \infty$, with $Q_\nu(t) = O(t^{-\nu-1}) = O(t^{-s})$. The integral representation

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

converges for $t > 1$ and $\mathrm{Re}(s) > 0$, is holomorphic in s, and one checks by differentiating under the integral that it solves Legendre's equation of degree $s-1$; it is the decaying solution. Hence $g_s = -2Q_{s-1}(1+\frac{1}{2}u)$ solves $(\Delta_z - s(s-1))g_s = 0$ off the diagonal. The normalizing factor -2 is fixed by the singularity computed next. □

5.3 The logarithmic singularity

Diagonal singularity. As $z \to w$, equivalently $t \to 1^+$,

$$g_s(z,w) = \log u(z,w) + O(1) = \log\frac{|z-w|^2}{y\,y'} + O(1),$$

so g_s has the logarithmic singularity required of a Green function, with $\frac{1}{\pi i}\partial\overline{\partial} g_s = \delta_{\mathrm{diag}} + (\mathrm{smooth})$.

Proof. The Legendre function of the second kind has the classical logarithmic behavior near $t = 1$,

$$Q_{s-1}(t) = -\tfrac{1}{2}\log\frac{t-1}{2} + \big(\psi(s) + \gamma\big) + o(1) \quad (t \to 1^+),$$

where ψ is the digamma function and γ Euler's constant; this follows from the hypergeometric expansion $Q_\nu(t) = \frac{1}{2}\log\frac{t+1}{t-1}\,P_\nu(t) - W_\nu(t)$ with W_ν analytic at $t = 1$ and $P_\nu(1) = 1$. Now $t - 1 = \frac{1}{2}u$, so

$$-2Q_{s-1}(t) = \log\frac{t-1}{2} - 2(\psi(s)+\gamma) + o(1) = \log u - \log 4 - 2(\psi(s)+\gamma) + o(1).$$

Thus $g_s = \log u + O(1)$ as claimed. Since $\log u = \log|z-w|^2 - \log(yy') = 2\log|z-w| + (\mathrm{smooth})$, the current $\frac{1}{\pi i}\partial\overline{\partial} g_s$ has the delta along the diagonal with the correct normalization, plus a smooth remainder coming from the $-s(s-1)g_s$ term and the $\log(yy')$ factor. □

5.4 Automorphizing over $\Gamma_0(N)$

Convergence of the automorphic Green function. For $\mathrm{Re}(s) > 1$ and $z \neq \gamma w$ for all γ, the series

$$G_s(z,w) = \sum_{\gamma \in \Gamma_0(N)} g_s(z, \gamma w)$$

converges absolutely and locally uniformly, defines a $\Gamma_0(N)$-invariant function of z and of w with a logarithmic singularity along the diagonal of $X_0(N)(\mathbb{C})$, and satisfies $(\Delta_z - s(s-1))G_s(\cdot, w) = 0$ away from the $\Gamma_0(N)$-orbit of w.

Proof. By Section 5.2, $Q_{s-1}(t) = O(t^{-s})$ as $t \to \infty$, so $g_s(z,\gamma w) = O\big(u(z,\gamma w)^{-s}\big) = O\big(e^{-s\,d(z,\gamma w)}\big)$. The number of $\gamma \in \Gamma_0(N)$ with $d(z,\gamma w) \le R$ grows like e^R (the hyperbolic lattice-point count, since $\Gamma_0(N)$ is a lattice and balls have exponentially growing volume $\sim e^R$). Therefore

$$\sum_{\gamma} e^{-s\,d(z,\gamma w)} \;\ll\; \int_0^\infty e^{-sR}\,d\big(e^R\big) = \int_0^\infty e^{(1-s)R}\,dR < \infty \quad \text{for } \mathrm{Re}(s) > 1,$$

so the series converges absolutely and locally uniformly. Invariance in z, w follows by reindexing, and the differential equation and single log singularity (from $\gamma = 1$) follow termwise. The physical Green function of Section 3.5 is recovered by removing the constant-eigenfunction contribution, which produces a simple pole at $s = 1$, and taking the finite part. □

5.5 Evaluation at Heegner points and the archimedean matching

The archimedean contribution to $\langle c, T_m c\rangle$ at a place of H over ∞ is, by Section 3.5 and the automorphic form of the Green function,

$$\langle c, T_m c\rangle_\infty = \sum_{\sigma,\tau}\sum_{\gamma} g_s\big(\tau_{\mathfrak{a}}^{\sigma}, \gamma\,\tau_{\mathfrak{b}}^{\tau}\big)\Big|_{\deg\gamma=m} = -2\sum_{\sigma,\tau}\sum_{\gamma} Q_{s-1}\big(1+\tfrac{1}{2}u(\tau_{\mathfrak{a}}^{\sigma}, \gamma\tau_{\mathfrak{b}}^{\tau})\big)\Big|_{\mathrm{c}}$$

a sum of Legendre functions at the point-pair invariants of CM points related by degree-m maps. The analytic side produces the same object.

Archimedean matching. The archimedean Fourier coefficient of the derivative of the Rankin–Selberg Eisenstein kernel at the center is, up to elementary Gamma factors, the same sum of Legendre functions $Q_{s-1}\big(1+\frac{1}{2}u\big)$ at the CM point-pair invariants. Hence the archimedean local height equals the archimedean analytic local term.

Proof. Compute the x-Fourier expansion of the free Green function and compare with the Bessel coefficient of Section 6.3. Fixing $w = iy'$ and $z = x + iy$, so $1 + \frac{1}{2}u(z,w) = (x^2+y^2+y'^2)/(2yy')$, the m-th Fourier coefficient of $g_s = -2Q_{s-1}(1+\frac{1}{2}u)$ in x is

$$\widehat{g}_s(m; y, y') = -2\int_{-\infty}^{\infty} Q_{s-1}\!\left(\frac{x^2+y^2+y'^2}{2yy'}\right) e^{-2\pi i m x}\,dx.$$

This has a classical closed form in modified Bessel functions: for $m \neq 0$, with $y_> = \max(y,y')$, $y_< = \min(y,y')$,

$$\widehat{g}_s(m; y, y') = -4\pi\sqrt{yy'}\; I_{s-\frac{1}{2}}\big(2\pi|m|y_<\big)\, K_{s-\frac{1}{2}}\big(2\pi|m|y_>\big),$$

the product of the two independent solutions I_ν (regular at 0) and K_ν (decaying at ∞) of the modified Bessel equation of order $\nu = s - \frac{1}{2}$. This is the standard evaluation of the resolvent kernel of the hyperbolic Laplacian by separation of variables: for fixed m, both g_s and its coefficient satisfy, in y, the ordinary differential equation obtained from $(\Delta_z - s(s-1))g_s = 0$ by inserting $e^{2\pi i m x}$, namely

$$-y^2\partial_y^2\widehat{g}_s + 4\pi^2 m^2 y^2 \widehat{g}_s = s(s-1)\widehat{g}_s,$$

whose two solutions are $\sqrt{y}\,I_{s-\frac{1}{2}}(2\pi|m|y)$ and $\sqrt{y}\,K_{s-\frac{1}{2}}(2\pi|m|y)$; the Green-function boundary conditions (regularity as $y \to y'^{-}$ from below, decay as $y \to \infty$) select the product $I_{s-\frac{1}{2}}(y_<)K_{s-\frac{1}{2}}(y_>)$, and the Wronskian $I_\nu(t)K_\nu'(t) - I_\nu'(t)K_\nu(t) = -1/t$ fixes the constant.

Now compare with the Eisenstein coefficient from Section 6.3: the m-th non-constant Fourier coefficient of $E(z,s)$ is proportional to $\sqrt{y}\,K_{s-\frac{1}{2}}(2\pi|m|y)$ times the divisor-sum $\sigma_{1-2s}(|m|)$. The archimedean part of the analytic local

term at index m is therefore the value at the CM point $y = \mathrm{Im}(\tau_{\mathfrak{a}})$ of exactly the same Whittaker factor $K_{s-\frac{1}{2}}(2\pi|m|y)$ that appears in $\widehat{g}_s(m; y, y')$.
Differentiating in s at the center $s_0 = 1$ commutes with all of this and produces $\partial_s K_{s-\frac{1}{2}}\big|_{s_0}$ on both sides. Hence, term by term in the Fourier index m,

$$\big(\text{archimedean height coefficient at } m\big) = \widehat{g}_s(m;\,\cdot\,)\Big|_{\mathrm{CM}} = \big(\text{archimedean analy}$$

and summing over m gives the equality of the archimedean local height with the archimedean analytic local term. The Legendre and Bessel forms are the same kernel written in the distance and Fourier variables. □

Thus, for every m and every place v, the local height $\langle c, T_m c\rangle_v$ equals a specific local coefficient of an analytic kernel. It remains to build that kernel, organize the local identities into an identity of modular forms, and extract the newform.

---

6. The analytic side: Rankin–Selberg integrals and the Eisenstein kernel

This section builds the object whose central derivative is $L'(E/K, 1)$: the unfolding identity, the Eisenstein Fourier expansion, the L-function, and the derivative at the center.

6.1 The newform and the theta series

Let $f = \sum_{n\geq 1} a_n q^n \in S_2(\Gamma_0(N))$ be the normalized newform of E. For an ideal class $A \in \mathrm{Cl}(\mathcal{O}_K)$ define the weight-one theta series

$$\theta_A(z) = \frac{1}{2u_K} + \sum_{n\geq 1} r_A(n)\, q^n, \qquad r_A(n) = \#\{\mathfrak{b} \subseteq \mathcal{O}_K \text{ integral} : [\mathfrak{b}] = A,\ \mathrm{N}\mathfrak{b} =$$

where $u_K = \frac{1}{2}\#\mathcal{O}_K^{\times}$. Summing over classes gives $r_K(n) = \sum_A r_A(n)$, the number of integral ideals of norm n, with Dirichlet series $\zeta_K(s) = \zeta(s)L(\chi_D, s)$. The Rankin–Selberg convolution of f against these theta series recovers $L(E/K, s)$, which the unfolding below makes precise.

6.2 The unfolding identity

Unfolding. Let $\Gamma = \Gamma_0(N)$, let $\Gamma_\infty = \{\pm \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)\}$ be the stabilizer of the cusp ∞, and let

$$E(z, s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \mathrm{Im}(\gamma z)^s, \qquad \mathrm{Re}(s) > 1,$$

be the weight-zero real-analytic Eisenstein series. Let F be a Γ-invariant function of rapid decay at the cusps, with Fourier expansion $F(z) = \sum_{n\in\mathbb{Z}} A_n(y)e^{2\pi inx}$ at ∞. Then

$$\int_{\Gamma\backslash\mathfrak{H}} F(z)E(z,s)\,d\mu(z) = \int_0^\infty A_0(y)\,y^{s-2}\,dy.$$

Proof. For $\mathrm{Re}(s) > 1$ the Eisenstein series converges absolutely and F decays rapidly, so we may interchange summation and integration:

$$\int_{\Gamma\backslash\mathfrak{H}} F(z) \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^s\,d\mu(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash\mathfrak{H}} F(z)\,\mathrm{Im}(\gamma z)^s\,d\mu(z).$$

In the γ-th term substitute $z \mapsto \gamma^{-1}z$. Since $F(\gamma^{-1}z) = F(z)$ by Γ-invariance, $d\mu$ is $\mathrm{SL}_2(\mathbb{R})$-invariant, and $\mathrm{Im}(\gamma\gamma^{-1}z)^s = \mathrm{Im}(z)^s = y^s$, the term becomes

$$\int_{\gamma(\Gamma\backslash\mathfrak{H})} F(z)\,y^s\,d\mu(z).$$

As γ ranges over a set of representatives for $\Gamma_\infty\backslash\Gamma$, the translated fundamental domains $\gamma(\Gamma\backslash\mathfrak{H})$ tile a fundamental domain for $\Gamma_\infty\backslash\mathfrak{H}$, without overlap and covering everything, because the coset space $\Gamma_\infty\backslash\Gamma$ reindexes exactly the Γ_∞-translates. Hence the sum of integrals is a single integral over $\Gamma_\infty\backslash\mathfrak{H}$:

$$\int_{\Gamma_\infty\backslash\mathfrak{H}} F(z)\,y^s\,d\mu(z).$$

A fundamental domain for $\Gamma_\infty\backslash\mathfrak{H}$ is the vertical strip $\{0 \le x < 1,\ y > 0\}$, since Γ_∞ is generated by $z \mapsto z+1$. Writing $d\mu = y^{-2}dx\,dy$ and integrating over $x \in [0,1)$ extracts the zeroth Fourier coefficient:

$$\int_0^\infty \Big(\int_0^1 F(z)\,dx\Big) y^s\,\frac{dy}{y^2} = \int_0^\infty A_0(y)\,y^{s-2}\,dy.$$

□

We assemble the L-function by recording the equal-weight Rankin–Selberg identity and then specializing; the mixed-weight case of f (weight 2) against θ_A (weight 1) folds the extra weight into the Eisenstein series.

Rankin–Selberg identity (equal weight). Let $f = \sum_{n\ge1} a_n q^n$ and $g = \sum_{n\ge1} b_n q^n$ be weight-k forms on $\Gamma_0(N)$ with f cuspidal, and let $E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \mathrm{Im}(\gamma z)^s$. Then for $\mathrm{Re}(s)$ large,

$$\int_{\Gamma_0(N)\backslash\mathfrak{H}} f(z)\,\overline{g(z)}\,E(z,s)\,y^k\,d\mu(z) = \frac{\Gamma(s+k-1)}{(4\pi)^{s+k-1}} \sum_{n\ge1} \frac{a_n\overline{b_n}}{n^{s+k-1}}.$$

Proof. The function $F(z) = f(z)\overline{g(z)}\,y^k$ is $\Gamma_0(N)$-invariant, because under γ the automorphy factors $j(\gamma,z)^k$ from f, $\overline{j(\gamma,z)}^k$ from $\overline{g}$, and $y^k \mapsto y^k/|j(\gamma,z)|^{2k}$ from the power of the imaginary part cancel; and F decays rapidly at every cusp because f is cuspidal. Its zeroth Fourier coefficient at ∞ is

$$A_0(y) = \int_0^1 f(z)\overline{g(z)}\, y^k \, dx = y^k \sum_{m,n\geq 1} a_n \overline{b_m}\, e^{-2\pi(n+m)y} \int_0^1 e^{2\pi i(n-m)x}\, dx = y^k \sum_{n}$$

since the x-integral is $\delta_{n=m}$. By the unfolding lemma of Section 6.2,

$$\int_{\Gamma_0(N)\backslash\mathfrak{H}} F(z)E(z,s)\, d\mu(z) = \int_0^\infty A_0(y)\, y^{s-2}\, dy = \sum_{n\geq 1} a_n \overline{b_n} \int_0^\infty e^{-4\pi n y}\, y^{s+k-2}$$

The inner integral is a Gamma integral: substituting $t = 4\pi n y$,

$$\int_0^\infty e^{-4\pi n y} y^{s+k-2}\, dy = \frac{\Gamma(s+k-1)}{(4\pi n)^{s+k-1}}.$$

Collecting the n-dependence gives the stated Dirichlet series times $\Gamma(s+k-1)(4\pi)^{-(s+k-1)}$. □

For the Gross–Zagier kernel the pairing is against $\theta_A(z)E_1(z,s)$ with E_1 a weight-one Eisenstein series; the same unfolding replaces $\overline{b_n}$ by $r_A(n)$, giving for one ideal class

$$\int_{\Gamma_0(N)\backslash\mathfrak{H}} f(z)\, \overline{\theta_A(z)}\, E_1(z,s)\, y\, d\mu(z) = \frac{\Gamma(s+1)}{(4\pi)^{s+1}} \sum_{n\geq 1} \frac{a_n\, r_A(n)}{n^{s+1}},$$

using $r_A(n) \in \mathbb{R}$. Summing over $A \in \mathrm{Cl}(\mathcal{O}_K)$ replaces $r_A(n)$ by $r_K(n) = \sum_A r_A(n)$, the number of integral ideals of norm n:

$$\sum_A \int_{\Gamma_0(N)\backslash\mathfrak{H}} f\, \overline{\theta_A}\, E_1(\cdot,s)\, y\, d\mu = \frac{\Gamma(s+1)}{(4\pi)^{s+1}} \sum_{n\geq 1} \frac{a_n\, r_K(n)}{n^{s+1}}.$$

The Dirichlet series factors as an Euler product because a_n is multiplicative (Hecke) and $r_K(n)$ is multiplicative with $\sum_n r_K(n) n^{-w} = \zeta_K(w) = \zeta(w)L(\chi_D, w)$; the convolution

$$\sum_{n\geq 1} \frac{a_n r_K(n)}{n^w} = \frac{L(f,w)\, L(f,\chi_D,w)}{\zeta(2w - 2\cdot\frac{k-1}{2})} \cdot (\cdots) = \frac{L(E/K,w)}{\zeta_N(2w-2)}$$

is the base-changed L-function divided by a zeta factor from the diagonal overcount, and the denominator is the normalizing factor of E_1, so clearing it leaves a completed L-function. With the archimedean Gamma factor, the completed function is

$$\Lambda(E/K,s) = \left(\frac{N|D|}{4\pi^2}\right)^s \Gamma(s)^2\, L(E/K,s),$$

the two Gamma factors coming from the two copies $\Gamma_{\mathbb{C}}(s) = (2\pi)^{-s}\Gamma(s)$ at the complex place of K, and it satisfies the functional equation

$$\Lambda(E/K,s) = w(E/K)\, \Lambda(E/K, 2-s) = -\Lambda(E/K, 2-s)$$

by Section 1.3, so $s_0 = 1$ is the center and $\Lambda(E/K,1) = 0$.

6.3 Fourier expansion of the Eisenstein series

The Fourier expansion of $E(z,s)$ comes by Poisson summation. We do the level-one computation in full; the level-N case differs only by finitely many Euler factors and the choice of cusp.

Fourier expansion of $E(z,s)$. For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$,

$$E(z,s) = y^s + \phi(s)\, y^{1-s} + \frac{2\pi^s \sqrt{y}}{\Gamma(s)\zeta(2s)} \sum_{n \neq 0} |n|^{s-\frac{1}{2}} \sigma_{1-2s}(|n|)\, K_{s-\frac{1}{2}}(2\pi|n|y)\, e^{2\pi i n x},$$

where $\sigma_\nu(n) = \sum_{d \mid n} d^\nu$, K_ν is the modified Bessel function, and the constant term coefficient is

$$\phi(s) = \sqrt{\pi}\, \frac{\Gamma(s - \tfrac{1}{2})}{\Gamma(s)}\, \frac{\zeta(2s-1)}{\zeta(2s)}.$$

Proof. Coset representatives for $\Gamma_\infty \backslash \Gamma$ are indexed by the bottom rows (c,d) with $\gcd(c,d) = 1$ up to sign, and $\mathrm{Im}(\gamma z) = y/|cz+d|^2$. Hence

$$E(z,s) = \tfrac{1}{2} \sum_{\gcd(c,d)=1} \frac{y^s}{|cz+d|^{2s}}.$$

Separate the terms with $c = 0$ (forcing $d = \pm 1$), which contribute y^s, from the terms with $c \neq 0$. For $c \neq 0$ group by $c \geq 1$ (using the $\pm$ symmetry) and remove the gcd condition by Möbius, writing every pair (c,d) with $c \geq 1$ as $g \cdot (c_0, d_0)$ with $\gcd(c_0,d_0) = 1$; this yields the factor $\zeta(2s)^{-1}$ after summing over the common divisor. Concretely,

$$E(z,s) = y^s + \frac{1}{\zeta(2s)} \sum_{c \geq 1} \sum_{d \in \mathbb{Z}} \frac{y^s}{|cz+d|^{2s}} \cdot \tfrac{1}{2} \cdot (\text{normalization}),$$

where the inner sum is now over all $d \in \mathbb{Z}$. Fix $c \geq 1$ and apply Poisson summation to $d \mapsto \Phi(d)$, $\Phi(t) = \dfrac{y^s}{\big((cx+t)^2 + c^2 y^2\big)^s}$:

$$\sum_{d \in \mathbb{Z}} \Phi(d) = \sum_{n \in \mathbb{Z}} \widehat{\Phi}(n), \qquad \widehat{\Phi}(n) = \int_{-\infty}^{\infty} \Phi(t) e^{-2\pi i n t}\, dt.$$

Substituting $t \mapsto t - cx$ and then $t \mapsto cy\, t$,

$$\widehat{\Phi}(n) = e^{2\pi i n c x}\, y^s \int_{-\infty}^{\infty} \frac{e^{-2\pi i n t}}{(t^2 + c^2 y^2)^s}\, dt = e^{2\pi i n c x}\, \frac{y^{1-s}}{c^{2s-1}} \int_{-\infty}^{\infty} \frac{e^{-2\pi i n c y t}}{(t^2+1)^s}\, dt.$$

The last integral is a classical Bessel integral: for $n = 0$,

$$\int_{-\infty}^{\infty} \frac{dt}{(t^2+1)^s} = \sqrt{\pi}\, \frac{\Gamma(s - \tfrac{1}{2})}{\Gamma(s)},$$

and for $n \neq 0$,

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi i\alpha t}}{(t^2+1)^s}\,dt = \frac{2\pi^s|\alpha|^{s-\frac12}}{\Gamma(s)}\,K_{s-\frac12}(2\pi|\alpha|), \qquad \alpha = ncy.$$

Summing $\widehat{\Phi}(n)$ over $c \geq 1$ and $n \in \mathbb{Z}$: the $n = 0$ terms give

$$\frac{1}{\zeta(2s)} \cdot \sqrt{\pi}\frac{\Gamma(s-\frac12)}{\Gamma(s)}\,y^{1-s}\sum_{c\geq 1} c^{1-2s} = \frac{\sqrt{\pi}\,\Gamma(s-\frac12)\,\zeta(2s-1)}{\Gamma(s)\,\zeta(2s)}\,y^{1-s} = \phi(s)y^{1-s},$$

which is the second constant-term piece. The $n \neq 0$ terms, after collecting the sum over c and n with nc fixed (the sum over divisors producing σ_{1-2s}), give

$$\frac{2\pi^s\sqrt{y}}{\Gamma(s)\zeta(2s)}\sum_{m\neq 0}|m|^{s-\frac12}\sigma_{1-2s}(|m|)\,K_{s-\frac12}(2\pi|m|y)\,e^{2\pi imx}.$$

Assembling the three pieces yields the stated expansion. □

For $\Gamma_0(N)$ and the weight required by Gross–Zagier (weight one, with the theta series absorbing the second weight, and a possible nebentypus), the same Poisson computation applies cusp by cusp; the constant term acquires the level-N scattering coefficients and the non-constant coefficients acquire the local factors at $p \mid N$, but the archimedean Bessel and the shape of the expansion are unchanged. The key structural output survives: the non-constant Fourier coefficients are products of a divisor-sum (the finite/nonarchimedean local data) and a Whittaker function (the archimedean local data).

6.4 Differentiating at the center

Since the root number is -1, the completed L-function vanishes at s_0, so the derivative is the informative quantity. Differentiating the integral representation moves ∂_s onto the Eisenstein series:

$$\frac{d}{ds}\Big|_{s_0}\sum_A \langle f, \theta_A E(\cdot,s)\rangle = \Big\langle f,\ \Theta\cdot\partial_s E(\cdot,s)\big|_{s_0}\Big\rangle \ \propto\ L'(f/K,s_0) = L'(E/K,1),$$

where $\Theta = \sum_A \theta_A$. The differentiated Eisenstein series $\partial_s E|_{s_0}$ is the analytic kernel. Its Fourier coefficients are the s-derivatives of the coefficients in Section 6.3:

$$\partial_s\big(y^s\big) = y^s\log y, \qquad \partial_s\big(\text{divisor-sum}\cdot K_{s-\frac12}(2\pi|m|y)\big) = (\partial_s\text{divisor-sum})K + \text{d}$$

Two features appear. The factor $\log y$, and the logarithmic derivatives $\partial_s(m^s) = m^s\log m$ in the divisor-sums, are the analytic counterpart of the intersection multiplicities ord_v and lifting lengths n of Section 4: differentiating an Euler factor produces the log matching the arithmetic multiplicity. The factor

$\partial_s K_{s-\frac{1}{2}}$ is the archimedean Whittaker derivative, equal by Section 5.5 to the Green function $Q_{s_0-1}(1+\frac{1}{2}u)$. Hence the Fourier coefficients of $\partial_s E|_{s_0}$ are the total local analytic terms, matching the local heights of Sections 4 and 5.

---

7. Holomorphic projection and modularity of the height series

The height generating series is holomorphic, but the analytic kernel $\partial_s E|_{s_0}$ is real-analytic. Holomorphic projection, which we prove through Poincaré series, bridges the two.

7.1 Poincaré series and their Petersson pairings

For weight $k \geq 2$, level N, index $m \geq 1$, and $\mathrm{Re}(s)$ large, define the Poincaré series

$$P_{m,s}(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \big(\mathrm{Im}\,\gamma z\big)^s \, e^{2\pi i m \gamma z} \, j(\gamma, z)^{-k},$$

where $j(\gamma, z) = cz + d$. It converges for $\mathrm{Re}(s) > 1 - k/2$, transforms with weight k, and its Petersson pairing against any weight-k automorphic function unfolds.

Unfolding of the Poincaré pairing. For a weight-k automorphic function Φ of moderate growth with Fourier expansion $\Phi(z) = \sum_n c_n(y) e^{2\pi i n x}$,

$$\langle \Phi, P_{m,s} \rangle = \int_{\Gamma_0(N)\backslash \mathfrak{H}} \Phi(z) \, \overline{P_{m,s}(z)} \, y^k \, d\mu(z) = \int_0^\infty c_m(y) \, e^{-2\pi m y} \, y^{s+k-2} \, \frac{dy}{1} \cdot (\text{con}$$

more precisely $\langle \Phi, P_{m,s} \rangle = \int_0^\infty c_m(y) \, e^{-2\pi m y} \, y^{s+k} \, \dfrac{dy}{y^2}$ after inserting the weight factor.

Proof. Insert the definition of $P_{m,s}$, conjugate, and unfold exactly as in Section 6.2. The pairing becomes an integral over $\Gamma_\infty \backslash \mathfrak{H}$,

$$\langle \Phi, P_{m,s} \rangle = \int_{\Gamma_\infty \backslash \mathfrak{H}} \Phi(z) \, y^s \, e^{-2\pi i m \bar{z}} \, y^k \, d\mu(z),$$

using $\overline{e^{2\pi i m z}} = e^{-2\pi i m \bar{z}}$ and $\overline{j(\gamma,z)^{-k}}$ combining with the weight factor to leave the invariant measure. On the strip $\{0 \leq x < 1\}$ write $e^{-2\pi i m \bar{z}} = e^{-2\pi i m x} e^{-2\pi m y}$; integrating over $x \in [0,1)$ against $\Phi = \sum_n c_n(y) e^{2\pi i n x}$ extracts the $n = m$ term $c_m(y)$. What remains is

$$\int_0^\infty c_m(y) \, e^{-2\pi m y} \, y^{s+k} \, \frac{dy}{y^2},$$

as claimed. □

7.2 The holomorphic projection lemma

Holomorphic projection. Let Φ transform with weight $k \geq 2$ on $\Gamma_0(N)$, be of moderate growth at the cusps, and have Fourier expansion $\Phi(z) = \sum_n c_n(y) e^{2\pi i n x}$. Then there is a holomorphic cusp form $\pi_{\mathrm{hol}}(\Phi) = \sum_{m \geq 1} c_m^{\mathrm{hol}} q^m \in S_k(\Gamma_0(N))$ with

$$c_m^{\mathrm{hol}} = \frac{(4\pi m)^{k-1}}{\Gamma(k-1)} \int_0^\infty c_m(y)\, e^{-4\pi m y}\, y^{k-2}\, dy,$$

characterized by $\langle g, \Phi \rangle = \langle g, \pi_{\mathrm{hol}}(\Phi) \rangle$ for all $g \in S_k(\Gamma_0(N))$. For $k = 2$ the integral is defined by analytic continuation in an auxiliary parameter, and the resulting form is a genuine cusp form.

Proof. First take $\Phi = g \in S_k$ holomorphic, so $c_m(y) = a_m(g) e^{-2\pi m y}$. By Section 7.1,

$$\langle g, P_{m,s} \rangle = \int_0^\infty a_m(g)\, e^{-2\pi m y}\, e^{-2\pi m y}\, y^{s+k-2}\, dy = a_m(g) \int_0^\infty e^{-4\pi m y} y^{s+k-2}\, dy =$$

Setting $s = 0$ gives $\langle g, P_{m,0} \rangle = a_m(g)\, \Gamma(k-1)/(4\pi m)^{k-1}$, so the holomorphic Poincaré series $P_m := P_{m,0}$ represents the m-th Fourier coefficient functional up to the constant $\Gamma(k-1)/(4\pi m)^{k-1}$. Since Fourier coefficients determine cusp forms, the P_m span S_k.

Now for general Φ, Section 7.1 at $s = 0$ gives

$$\langle \Phi, P_m \rangle = \int_0^\infty c_m(y)\, e^{-2\pi m y}\, y^{k-2}\, dy.$$

Define $\pi_{\mathrm{hol}}(\Phi) = \sum_{m \geq 1} c_m^{\mathrm{hol}} q^m$ with c_m^{hol} as displayed; then by the holomorphic computation above,

$$\langle \pi_{\mathrm{hol}}(\Phi), P_m \rangle = c_m^{\mathrm{hol}}\, \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} = \int_0^\infty c_m(y)\, e^{-2\pi m y}\, y^{k-2}\, dy \;=\; \langle \Phi, P_m \rangle,$$

where the normalizing constant is chosen exactly so that the two pairings against every P_m agree; since the P_m span S_k, this forces $\langle g, \pi_{\mathrm{hol}}(\Phi) \rangle = \langle g, \Phi \rangle$ for all $g \in S_k$. Uniqueness of $\pi_{\mathrm{hol}}(\Phi)$ inside S_k follows because a cusp form orthogonal to all of S_k is zero. For $k = 2$ the integral $\int_0^\infty c_m(y) e^{-4\pi m y} y^0 dy$ sits at the edge of convergence when $c_m(y)$ grows; one restores convergence by inserting y^s, computing for $\mathrm{Re}(s) > 0$, and continuing to $s = 0$, which is legitimate because the divergent boundary contribution is a constant (the non-cuspidal part) orthogonal to S_2 and hence invisible to the pairing. The continued value defines c_m^{hol} and $\pi_{\mathrm{hol}}(\Phi) \in S_2$. □

7.3 Modularity of the height generating series

Applying holomorphic projection to the weight-two analytic kernel $\Phi(z, s_0) = \big(\Theta \cdot \partial_s E_1(\cdot, s)\big)\big|_{s_0}$, whose m-th Fourier coefficient $c_m(y)$ is the total (archimedean and nonarchimedean) analytic local term at index m before projection, produces a holomorphic weight-two cusp form

$$\Psi = \pi_{\mathrm{hol}}\big(\Phi(\cdot, s_0)\big) \in S_2(\Gamma_0(N)), \qquad a_m(\Psi) = \frac{4\pi m}{\Gamma(1)} \int_0^\infty c_m(y)\, e^{-4\pi m y}\, dy,$$

the $k = 2$ case of the holomorphic projection formula, whose m-th Fourier coefficient is the holomorphically projected total analytic local term for index m. On the geometric side define the height generating series

$$\mathcal{H}(z) = \sum_{m \geq 1} \langle c, T_m c \rangle\, q^m.$$

Modularity of the height series. The generating series $\mathcal{H}$ is a holomorphic cusp form of weight two on $\Gamma_0(N)$, and $\mathcal{H} = \Psi$.

Proof. By Sections 4 and 5 the local heights match the local analytic terms place by place: for each m,

$$\langle c, T_m c \rangle = \sum_v \langle c, T_m c \rangle_v = \Big(\text{holomorphically projected } m\text{-th coefficient of } \partial_s E$$

The holomorphic projection was needed precisely so that the archimedean and non-holomorphic contributions on the analytic side are replaced by the coefficient of a genuine weight-two form, and the matching identities of Sections 4 and 5 are exactly the statement that after this projection the coefficient equals the local height sum. Since $\mathcal{H}$ and Ψ have identical Fourier coefficients for all $m \geq 1$ and $\Psi \in S_2(\Gamma_0(N))$, the series $\mathcal{H}$ is that same cusp form. In particular $\mathcal{H}$ is modular of weight two; this is the analytic incarnation of the geometric fact, established independently and in sharper form by Gross, Kohnen, and Zagier, that Heegner divisor classes vary modularly. □

---

8. The comparison and extraction of the newform

Both $\mathcal{H}$ and Ψ lie in $S_2(\Gamma_0(N))$ and are equal. Pairing with f isolates the height on one side and $L'(E/K, 1)$ on the other, using Hecke self-adjointness (proved next) and multiplicity one (cited).

8.1 Hecke operators are self-adjoint for the height pairing

Self-adjointness of T_m. For $(m, N) = 1$, the Hecke operator T_m acting on $J_0(N)$ is self-adjoint for the Néron–Tate height pairing:

$$\langle T_m a, b\rangle = \langle a, T_m b\rangle, \qquad a, b \in J_0(N)(F).$$

Proof. The Néron–Tate pairing on $J = J_0(N)$ is $\langle a, b\rangle = \hat{h}_\Theta(a, b)$ for the symmetric ample class Θ giving the canonical principal polarization $\lambda_\Theta : J \xrightarrow{\sim} J^\vee$. For an endomorphism $\phi \in \mathrm{End}(J)$, the adjoint with respect to this pairing is the Rosati involution $\phi^\dagger = \lambda_\Theta^{-1} \circ \phi^\vee \circ \lambda_\Theta$, characterized by $\langle \phi a, b\rangle = \langle a, \phi^\dagger b\rangle$; this is a formal consequence of the functoriality of the height pairing under isogenies and the definition of $\phi^\vee$ as the dual morphism.

It remains to show $T_m^\dagger = T_m$ for $(m, N) = 1$. Realize T_m as the correspondence on $X_0(N)$ given by the curve $X_0(N, m)$ parametrizing cyclic m-isogenies $(A, C) \to (A', C')$, with the two projections π_1, π_2 to $X_0(N)$. The transpose correspondence T_m^t swaps π_1 and π_2, that is, it parametrizes the dual isogenies. For $(m, N) = 1$ the dual of a cyclic m-isogeny is again a cyclic m-isogeny, and the involution sending an isogeny to its dual is implemented on $X_0(N, m)$ by an automorphism compatible with the level-N structure, so $T_m^t = T_m$ as correspondences. Under the Albanese functoriality that defines the action on $J_0(N)$, the transpose of a correspondence induces the Rosati adjoint of the induced endomorphism, since the canonical polarization is defined by the theta divisor and correspondences act compatibly with it. Therefore $T_m^\dagger = T_m$, giving self-adjointness. □

8.2 Extracting the height and the derivative

On the geometric side, decompose $S_2(\Gamma_0(N))$ into Hecke eigenforms; f spans a one-dimensional eigenline. Let $c^{(f)}$ be the f-isotypic projection of c, with $T_m c^{(f)} = a_m c^{(f)}$. Self-adjointness of T_m (Section 8.1) gives, for every m,

$$\langle c, T_m c\rangle = \langle T_m c, c\rangle, \qquad \text{and} \qquad \langle c, T_m c\rangle^{(f)} = \langle c^{(f)}, T_m c^{(f)}\rangle = a_m \langle c^{(f)}, c^{(f)}\rangle,$$

where the superscript (f) denotes the component of the sequence $m \mapsto \langle c, T_m c\rangle$ landing in the f-eigensystem, extracted because eigenforms with distinct systems are height-orthogonal. Hence the f-isotypic part of the generating series is

$$\mathcal{H}^{(f)} = \sum_m a_m \langle c^{(f)}, c^{(f)}\rangle\, q^m = \langle c^{(f)}, c^{(f)}\rangle\, f,$$

a scalar multiple of f, and by Petersson orthogonality

$$\langle \mathcal{H}, f\rangle_{\mathrm{Pet}} = \langle c^{(f)}, c^{(f)}\rangle \, \|f\|^2.$$

It remains to relate $\langle c^{(f)}, c^{(f)}\rangle$ to $\hat{h}(P_K)$. Since $P_K = \varphi_*(y_K)$ and the trace y_K has the same f-isotypic projection as c up to the Galois averaging already built into P_K, the projection formula for the pushforward φ_* and its dual φ^* gives

$$\hat{h}(P_K) = \langle \varphi_* y_K, \varphi_* y_K\rangle_E = \langle y_K, \varphi^*\varphi_* y_K\rangle_J.$$

The composite $\varphi^*\varphi_*$ acts on the f-isotypic line of $J_0(N)$ as multiplication by $\deg\varphi$, because φ is the optimal quotient onto E up to the Manin constant and $\varphi \circ \varphi^\vee = [\deg\varphi]$ on E. Therefore

$$\hat{h}(P_K) = \deg\varphi \cdot \langle y_K^{(f)}, y_K^{(f)}\rangle_J = \deg\varphi \cdot \langle c^{(f)}, c^{(f)}\rangle \cdot u_K^{-2} \cdot (\text{power of } 2),$$

the factor u_K^{-2} from the relation between the full Heegner divisor y_K and the single class c under the trace. Combining,

$$\langle \mathcal{H}, f\rangle_{\mathrm{Pet}} = \frac{\hat{h}(P_K)}{\deg\varphi} \cdot \|f\|^2 \cdot u_K^2 \cdot (\text{power of } 2).$$

On the analytic side, since Ψ is the holomorphic projection of the differentiated kernel and holomorphic projection preserves Petersson pairings against cusp forms,

$$\langle \Psi, f\rangle_{\mathrm{Pet}} = \Big\langle f,\ \partial_s\big(\Theta E_1(\cdot, s)\big)\big|_{s_0}\Big\rangle = \frac{d}{ds}\Big|_{s_0}\Big\langle f,\ \Theta E_1(\cdot, s)\Big\rangle = \frac{d}{ds}\Big|_{s_0} \frac{\Gamma(s+1)}{(4\pi)^{s+1}} \frac{L(I$$

At the center s_0 the completed L-function vanishes, so the s-derivative lands entirely on $L(E/K, \cdot)$, giving a nonzero multiple of $L'(E/K, 1)$ with the explicit archimedean constant $\Gamma(2)(4\pi)^{-2} = (4\pi)^{-2}$ times the value of the zeta normalization at the center.

The Gross–Zagier identity. There is an explicit nonzero constant $C(E, K, \varphi)$ with

$$\hat{h}(P_K) = C(E, K, \varphi)\, L'(E/K, 1).$$

Proof. Since $\mathcal{H} = \Psi$ by Section 7.3, their Petersson pairings against f coincide, $\langle \mathcal{H}, f\rangle_{\mathrm{Pet}} = \langle \Psi, f\rangle_{\mathrm{Pet}}$. Substituting the two explicit evaluations from Section 8.2,

$$\frac{\hat{h}(P_K)}{\deg\varphi} \cdot \|f\|^2 \cdot u_K^2 \cdot 2^a = \frac{d}{ds}\Big|_{s_0} \frac{\Gamma(s+1)}{(4\pi)^{s+1}} \frac{L(E/K, s+1)}{\zeta_N(2s)} = \kappa \cdot L'(E/K, 1),$$

where 2^a is the power of two from the trace normalization and κ is the nonzero archimedean-and-zeta constant obtained by differentiating the completed factor at the center (using $L(E/K, 1) = 0$, so only the derivative of L survives).

Multiplicity one for f (cited) makes the f-isotypic line one-dimensional, so this is a genuine scalar identity rather than a proportionality up to unknown vectors. Solving,

$$\hat{h}(P_K) = C(E,K,\varphi)\, L'(E/K,1), \qquad C(E,K,\varphi) = \frac{\kappa\, \deg\varphi}{2^a\, u_K^2\, \|f\|^2} > 0.$$

Every factor is positive: $\deg\varphi \geq 1$, $u_K \geq 1$, $\|f\|^2 > 0$, and $\kappa > 0$ is the product of the positive archimedean Gamma factor and the positive value of the zeta normalization at the center. Hence $C(E,K,\varphi) > 0$, giving the equivalence $L'(E/K,1) \neq 0 \iff \hat{h}(P_K) \neq 0 \iff P_K$ of infinite order. Matching κ and 2^a against the explicit constants of Sections 6 and 8 yields the closed form recorded in Section 9. □

---

9. The precise formula and its constant

For definiteness we record the identity with its constant assembled. Let f be the newform of E with Petersson norm

$$\|f\|^2 = \int_{\Gamma_0(N)\backslash\mathfrak{H}} |f(z)|^2\, y^2\, d\mu(z),$$

let $D = \mathrm{disc}(K)$, let $u_K = \frac{1}{2}\#\mathcal{O}_K^\times$, and let $\varphi : X_0(N) \to E$ be the parametrization with $\varphi^*\omega_E = c_\varphi \cdot 2\pi i f(\tau)\, d\tau$. The Gross–Zagier identity in its Petersson-normalized form is

$$L'(f/K,1) = \frac{8\pi^2\, \|f\|^2}{u_K^2 \sqrt{|D|}} \cdot \frac{\hat{h}(P_K)}{\deg\varphi} \cdot c_\varphi^{-2},$$

so that

$$\hat{h}(P_K) = \frac{u_K^2 \sqrt{|D|}\, \deg\varphi\, c_\varphi^2}{8\pi^2\, \|f\|^2}\, L'(f/K,1).$$

The Manin constant and exact power of two depend on the normalization of φ; the structural content is the strict proportionality with a positive constant, which feeds the regulator into the rank-one Birch–Swinnerton-Dyer formula.

---

10. Kolyvagin’s Euler system and the rank-one BSD consequence

Gross–Zagier gives one inequality: $L'(E/K,1) \neq 0$ forces $\mathrm{rank}\, E(K) \geq 1$. Kolyvagin’s Euler system of Heegner points pins the rank at one and bounds Sha, citing only the two structural inputs of the status note.

10.1 Heegner points over ring class fields

For a squarefree product $n = \ell_1 \cdots \ell_r$ of Kolyvagin primes (rational primes inert in K with $\ell \nmid N$ and $a_\ell \equiv \ell + 1 \equiv 0 \pmod{p^M}$ for the working modulus p^M), CM by the order $\mathcal{O}_n = \mathbb{Z} + n\mathcal{O}_K$ gives Heegner points $P_n \in E(K_n)$ over the ring class field K_n of conductor n, with $\mathrm{Gal}(K_n/K) \cong \mathrm{Cl}(\mathcal{O}_n)$, exactly as in Section 2. Two relations hold.

Euler system relations. For a Kolyvagin prime $\ell \nmid n$, with σ_ℓ a generator of the cyclic group $\mathrm{Gal}(K_{n\ell}/K_n)$ of order $\ell + 1$: 1. norm compatibility: $\mathrm{Tr}_{K_{n\ell}/K_n}(P_{n\ell}) = a_\ell\, P_n$; 2. congruence: $P_{n\ell} \equiv \mathrm{Frob}_\ell(P_n) \pmod{\lambda}$ for a prime $\lambda \mid \ell$ of $K_{n\ell}$, where Frob_ℓ is the Frobenius at ℓ.

Proof. Both relations come from the Eichler–Shimura congruence and the moduli interpretation of Hecke operators. The trace $\mathrm{Tr}_{K_{n\ell}/K_n}$ sums the Galois conjugates of $P_{n\ell}$ over $\mathrm{Gal}(K_{n\ell}/K_n)$, which by Shimura reciprocity for the order $\mathcal{O}_{n\ell}$ are the $\ell + 1$ Heegner points of conductor $n\ell$ lying above P_n; these are exactly the ℓ-isogenous CM points counted by T_ℓ, so their sum is T_ℓ applied to P_n, which equals $a_\ell P_n$ since P_n lies in the f-isotypic part where T_ℓ acts by a_ℓ. Relation (2) is the reduction of this identity modulo a prime above ℓ: by Eichler–Shimura, $T_\ell \equiv \mathrm{Frob}_\ell + \ell\langle \ell \rangle \,\mathrm{Frob}_\ell^{-1} \pmod{\ell}$, and evaluating the CM points modulo λ collapses the conductor-$n\ell$ point to the Frobenius image of the conductor-n point, giving the stated congruence. □

10.2 Derived classes

Kolyvagin's derivative operators turn this norm-compatible family into cohomology classes. For each $\ell \mid n$ set

$$D_\ell = \sum_{i=1}^{\ell} i\, \sigma_\ell^i \in \mathbb{Z}[\mathrm{Gal}(K_{n\ell}/K_n)], \qquad D_n = \prod_{\ell \mid n} D_\ell.$$

The operator D_ℓ satisfies the telescoping identity $(\sigma_\ell - 1)D_\ell = (\ell + 1) - \mathrm{Tr}_\ell$, where $\mathrm{Tr}_\ell = \sum_i \sigma_\ell^i$.

Construction of the Kolyvagin classes. The point $D_n P_n \in E(K_n)/p^M$ is fixed by $\mathrm{Gal}(K_n/K)$ modulo p^M, and it defines a canonical class

$$c(n) \in H^1\big(K, E[p^M]\big),$$

whose local behavior is unramified outside n and precisely controlled at the primes dividing n.

Proof. We check $\mathrm{Gal}(K_n/K)$-invariance of $D_n P_n$ modulo p^M prime by prime. For $\ell \mid n$,

$$(\sigma_\ell - 1) D_n P_n = D_{n/\ell}\big((\ell+1) - \mathrm{Tr}_\ell\big) P_n.$$

By norm compatibility, $\mathrm{Tr}_\ell\, P_n = a_\ell P_{n/\ell}$ (viewed in $E(K_n)$ after the natural inclusion), while $(\ell+1)P_n$ is the other term. Since ℓ is a Kolyvagin prime, $a_\ell \equiv 0$ and $\ell + 1 \equiv 0 \pmod{p^M}$, so both terms vanish modulo p^M. Hence $(\sigma_\ell - 1) D_n P_n \equiv 0 \pmod{p^M}$ for every $\ell \mid n$, and since these σ_ℓ generate $\mathrm{Gal}(K_n/K)$, the class $D_n P_n$ is $\mathrm{Gal}(K_n/K)$-invariant modulo p^M.

Now use the inflation-restriction sequence

$$0 \to H^1(K_n/K, E(K_n)[p^M]) \to H^1(K, E[p^M]) \xrightarrow{\mathrm{res}} H^1(K_n, E[p^M])^{\mathrm{Gal}(K_n/K)}.$$

The invariant point $D_n P_n$ gives a class in $E(K_n)/p^M$, which maps to a $\mathrm{Gal}(K_n/K)$-invariant element of $H^1(K_n, E[p^M])$ via the Kummer map; invariance and the vanishing of the obstruction group (which holds because $E[p^M]$ has no K_n/K-invariants under the working hypotheses, a consequence of the irreducibility of the mod-p representation) lift it uniquely to $c(n) \in H^1(K, E[p^M])$. The local conditions follow from the Euler system relations: at primes not dividing n the class is unramified because P_n is, and at $\ell \mid n$ the congruence relation (2) computes the ramified part in terms of the Frobenius image, giving the controlled local behavior. □

10.3 Bounding the Selmer group

Kolyvagin's theorem. If $P_K \in E(K)$ has infinite order, then $E(K) \otimes \mathbb{Q}$ is one-dimensional and $\mathrm{Sha}(E/K)[p^\infty]$ is finite.

Proof. The classes $c(n) \in H^1(K, E[p^M])$ satisfy the Euler system local conditions: $c(n)$ is unramified away from n, and at each $\ell \mid n$ its ramified component is nonzero and computes the localization of $c(n/\ell)$ at ℓ via relation (2).

Suppose for contradiction that the p^M-Selmer group $\mathrm{Sel}_{p^M}(E/K)$ has rank exceeding one. Choose a nonzero class $s \in \mathrm{Sel}_{p^M}(E/K)$ not proportional to the class of P_K. By the global duality reciprocity law, for any global class the sum of local invariants of the cup product vanishes:

$$\sum_{v} \mathrm{inv}_v\big(s \cup c(n)\big) = 0.$$

We choose the Kolyvagin prime ℓ by Chebotarev so that the localizations s_ℓ and $(P_K)_\ell$ are simultaneously controlled: since the mod-p representation on $E[p^M]$ has large image, there is a positive-density set of primes ℓ (inert in K, Kolyvagin) at which the Frobenius conjugacy class realizes any prescribed pair of local behaviors for s and for the Heegner class. For such ℓ, all local invariants

in the reciprocity sum vanish except the one at ℓ, which pairs the ramified part of $c(\ell)$ (nonzero, by the infinite order of P_K feeding relation (2)) against the localization s_ℓ. Vanishing of the total sum forces $s_\ell = 0$. Ranging over enough such ℓ forces s to be locally trivial everywhere, hence globally trivial by the injectivity of the localization map on the relevant Selmer quotient, contradicting $s \neq 0$.

Therefore $\mathrm{Sel}_{p^M}(E/K)$ has rank at most one for all M, so its projective limit, the p-adic Selmer group, has corank at most one. The exact sequence

$$0 \to E(K) \otimes \mathbb{Q}_p/\mathbb{Z}_p \to \mathrm{Sel}_{p^\infty}(E/K) \to \mathrm{Sha}(E/K)[p^\infty] \to 0$$

then gives $\operatorname{rank} E(K) \leq 1$; combined with the infinite order of P_K this forces $\operatorname{rank} E(K) = 1$, and it forces $\mathrm{Sha}(E/K)[p^\infty]$ to be finite because the corank-one Selmer group is exhausted by the free part. Running the argument at all p gives finiteness of $\mathrm{Sha}(E/K)$. □

10.4 Descent to $\mathbb{Q}$

Rank-one Birch–Swinnerton-Dyer over $\mathbb Q$. If $\operatorname{ord}_{s=1} L(E,s) = 1$, then $\operatorname{rank} E(\mathbb{Q}) = 1$ and $\mathrm{Sha}(E/\mathbb{Q})$ is finite.

Proof. Choose auxiliary imaginary quadratic K satisfying the Heegner hypothesis for N with $L(E,\chi_K,1) \neq 0$ (such K exist by nonvanishing theorems for quadratic twists). Then $\operatorname{ord}_{s=1} L(E/K,s) = 1$, so $L'(E/K,1) \neq 0$; by Gross–Zagier P_K has infinite order, and by Kolyvagin $\operatorname{rank} E(K) = 1$ with $\mathrm{Sha}(E/K)$ finite. Decomposing $E(K) \otimes \mathbb{Q}$ into χ_K-eigenspaces, the plus part is $E(\mathbb{Q}) \otimes \mathbb{Q}$ and the minus part $E^{(D)}(\mathbb{Q}) \otimes \mathbb{Q}$; since $L(E,\chi_K,1) \neq 0$, the twist has rank zero, placing all of the rank in the plus part. Hence $\operatorname{rank} E(\mathbb{Q}) = 1$ and $\mathrm{Sha}(E/\mathbb{Q})$ is finite. □

---

11. Comparing both sides

Both sides are assembled from the same input, the CM data of K on the modular curve, and matched place by place: at finite primes, intersection multiplicities of CM points are counts of quaternionic embeddings matched to derivatives of local Whittaker functions counting the same reduced-norm solutions; at the archimedean place, both the Green function and the differentiated Eisenstein coefficient are Q_{s-1} at the same point-pair invariants. We summarize this as follows:

Geometric side	Analytic side
CM points on $X_0(N)$	theta series attached to K
Heegner divisors	Fourier coefficients indexed by norms
Néron–Tate height	central derivative of $L(f/K, s)$
finite local intersections	derivatives of local Whittaker functions
optimal embeddings into quaternion orders	local representation densities
lifting lengths (Serre–Tate)	logarithmic derivatives of Euler factors
archimedean Green functions	differentiated Eisenstein coefficients

The theorem thus equates the arithmetic height of CM cycles with the central derivative of the automorphic object they generate, place by place.

---

12. Generalizations

Gross–Zagier is the first instance of a principle relating heights of algebraic cycles to derivatives of L-functions. The Gross–Kohnen–Zagier theorem shows that Heegner divisor classes of varying discriminant in $J_0(N) \otimes \mathbb{Q}$ are the Fourier coefficients of a weight-$3/2$ modular form, so Heegner points vary modularly. The Kudla program generalizes the geometric side to special cycles on orthogonal and unitary Shimura varieties, predicting that their generating series are modular and their arithmetic intersection pairings compute derivatives of Eisenstein series, with Gross–Zagier the one-dimensional case. The arithmetic Gan–Gross–Prasad program relates heights of diagonal cycles on products of Shimura varieties to central derivatives of Rankin–Selberg L-functions; the work of Yuan, Zhang, and Zhang gives a general Gross–Zagier formula for Shimura curves over totally real fields, reorganized through automorphic representations; and Waldspurger's formula, expressing central L-values through toric periods, sits at the analytic-rank-zero end. The unifying expectation is that central derivatives of L-functions are heights of special cycles.

Denouement

Let us summarize this spectacular proof in words. The Heegner hypothesis makes $w(E/K) = -1$ and $L(E/K, 1) = 0$. The ideal classes of K give CM elliptic curves, which the auxiliary ideal $\mathfrak{n}$ equips with cyclic N-structure to produce Heegner points on $X_0(N)$, defined over H and permuted by the class group. Summing them and pushing through φ gives $P_K \in E(K)$. Its Néron–Tate height decomposes into local terms: intersection numbers on regular models, computed by the deformation theory of CM points and the quaternionic counting at nonsplit primes, and archimedean Green-function values, computed through the Legendre resolvent of the hyperbolic Laplacian. The same CM data builds the theta series whose Rankin–Selberg convolution with f is $L(E/K, s)$; unfolding, differentiation at the center, and holomorphic projection turn its central derivative into a weight-two cusp form. Place by place, the local heights match the local analytic terms, so the two weight-two forms agree, and pairing with f through Hecke self-adjointness and multiplicity one yields $\hat{h}(P_K) = C(E, K, \varphi)L'(E/K, 1)$ with $C > 0$. Combined with Kolyvagin's Euler system, this gives the Birch–Swinnerton-Dyer conjecture in analytic rank at most one, and it is the blueprint for the principle that central derivatives of L-functions are heights of special cycles.

References

- Benedict H. Gross and Don B. Zagier, *Heegner points and derivatives of L-series*, Inventiones Mathematicae 84 (1986), 225–320.
- Benedict H. Gross, *Heegner points on $X_0(N)$*, in *Modular Forms* (Durham, 1983), Ellis Horwood, 1984.
- Benedict H. Gross, Winfried Kohnen, and Don B. Zagier, *Heegner points and derivatives of L-series II*, Mathematische Annalen 278 (1987), 497–562.
- Victor A. Kolyvagin, *Euler systems*, in *The Grothendieck Festschrift II*, Progress in Mathematics 87, Birkhäuser, 1990.
- Henri Darmon, *Rational Points on Modular Elliptic Curves*, CBMS 101, American Mathematical Society, 2004.
- Joseph H. Silverman, *The Arithmetic of Elliptic Curves*, GTM 106, Springer.
- Serge Lang, *Elliptic Functions*, GTM 112, Springer.
- David A. Cox, *Primes of the Form $x^2 + ny^2$*, Wiley.
- Daniel Bump, *Automorphic Forms and Representations*, Cambridge University Press.

- Henryk Iwaniec, *Spectral Methods of Automorphic Forms*, GSM 53, American Mathematical Society.
- Serre–Tate, deformation theory of abelian varieties, for the local deformation input.
- Xinyi Yuan, Shou-Wu Zhang, and Wei Zhang, *The Gross–Zagier Formula on Shimura Curves*, Annals of Mathematics Studies 184, Princeton University Press, 2013.
- Shou-Wu Zhang, *Gross–Zagier formula for* GL_2, Asian Journal of Mathematics 5 (2001), 183–290.