Modular curves and local heights

These are notes meant to explain the computations of local height pairings of Heegner points on modular curves, as done by Gross and Zagier. These computations comprise one side of the famous Gross–Zagier formula, which we state in §3. The other side of the formula (computing coefficients of Rankin L-series) is not a focus of these notes.

Contents

1 Modular curves I

In this section we will give descriptions of the smooth curves $Y_0(N), Y_1(N), X_0(N), X_1(N)$ living over $\mathbf{Z}[1/N]$. We will ignore the underlying stack-theoretic constructions for now, they will come to the forefront regardless in §2.

1.1 The curves $Y_0(N)$ and $Y_1(N)$

We recall the definitions of the modular curves $Y_0(N)$ and $Y_1(N)$. Consider the functors $F_0, F_1 : (\mathrm{Sch}/\mathbf{Z}[1/N])^{\mathrm{opp}} \to \mathrm{Set}$ given by

$$F_0(S) = \{\phi : E \to E'\}, \qquad F_1(S) = \{(E, P)\}$$

where E, E' are elliptic curves over S, $\phi : E \to E'$ is a cyclic S-isogeny of degree N (one whose kernel is étale locally isomorphic to $(\mathbf{Z}/N\mathbf{Z})_S$), and $P \in E(S)$ is a point which has exact order N in every geometric fiber. There is a surjective map of sheaves $F_1 \to F_0$ which sends (E, P) to the isogeny $\phi : E \to E/\langle P \rangle$. We can also view F_1 as the functor

$$F_1(S) = \{(\phi : E \to E', P)\}$$

where P is a generator of $\ker \phi$. Then the map $F_1 \to F_0$ is just forgetful.

The functor F_1 is representable by a scheme for $N \geq 4$, smooth of relative dimension 1 over $\operatorname{Spec} \mathbf{Z}[1/N]$. We omit the proof of representability, but it follows from the existence of the Tate normal form of an elliptic curve with a marked point of order ≥ 4 (possibly infinite). Assuming representability, we can prove:

Proposition 1.1. *$Y_1(N)$ is a smooth curve over $\mathbf{Z}[1/N]$.*

Proof. Let R be a local $\mathbf{Z}[1/N]$-algebra, I an ideal of square zero, and $\overline{R} = R/I$. Then $\overline{R}$ is again local. Let $(\overline{E}, \overline{P})$ an elliptic curve over $\overline{R}$. Now $\overline{E}$ has a Weierstrass model in $\mathbf{P}^2_{\overline{R}}$, lift it to a closed subscheme of $\mathbf{P}^2_R$ by lifting the coefficients. The discriminant is automatically a unit, and so this gives an elliptic curve E/R lifting $\overline{E}$, i.e. $E_{\overline{R}} \cong \overline{E}$.

Now $E[N]$ is finite étale over R, since N is invertible in R, so the natural map $E[N](R) \to \overline{E}[N](\overline{R})$ is an isomorphism and we can find a lift of $\overline{P}$.

We have established the smoothness of $Y_1(N)$; it remains to establish its relative dimension. For this we can work on the generic fiber $Y_1(N)_{\mathbf{Q}}$ (or really any fiber). The tangent space at (E, P) corresponds to flat first-order infinitesimal deformations of the pair (E, P). We claim that such deformations of (E, P) are the same as deformations of E; indeed given a deformation of E, the closure of P provides a compatible deformation of P. Deformations of E are parametrized by $H^1(E, T_E)$ which has dimension 1 over $\mathbf{Q}$. □

However, the functor F_0 is not representable, evidently the data it parametrizes always has automorphisms. It's coarse moduli space is again a smooth curve $Y_0(N)/\operatorname{Spec} \mathbf{Z}[1/N]$. It may be obtained as the quotient of $Y_1(N)$ by the action of $(\mathbf{Z}/N\mathbf{Z})^\times$ on the $\Gamma_1(N)$-level structure.

1.2 Compactifications

There are a few ways to write down compactifications of $Y_0(N)$ and $Y_1(N)$. One way, due to Deligne–Rapoport [2], is to consider an appropriate moduli functor of *generalized elliptic curves*.

Definition 1.2. A *generalized elliptic curve* over an algebraically closed field k will refer to a curve E/k such that

- E is reduced, connected, and has at-worst nodal singularities.
- E has arithmetic genus 1.
- E has trivial dualizing sheaf.
- E^{sm} is equipped with the structure of a group scheme and an action map $E^{\mathrm{sm}} \times E \to E$ which extends the group law on E^{sm} and acts transitively on the irreducible components of E.

A *generalized elliptic curve* over an arbitrary base S is the data of a flat, proper S-scheme E, an S-group structure on E^{sm} which extends to $E^{\mathrm{sm}} \times_S E \to E$, and such that all geometric fibers are generalized elliptic curves as above.

This definition captures all the key properties of an elliptic curve, except for smoothness.

Proposition 1.3. *A generalized elliptic curve over $k = \overline{k}$ is isomorphic either to an elliptic curve or to a Néron polygon.*

Definition 1.4. A *Néron polygon* over a scheme S is one obtained by taking n copies of $\mathbf{P}^1_S$, labeled $\ell_1, \ldots, \ell_n$, and gluing $0 \in \ell_i$ to $\infty \in \ell_{i+1}$ (indices taken modulo n).

We note a few properties of this definition. Let C_n denote a Néron n-gon. First, we see that C_1 is simply a nodal cubic. In general, an n-gon is a (singular) curve of genus 1 (compute its Euler characteristic). Second, we can easily compute that the (relative) dualizing sheaf is trivial. Indeed, $\omega_{\mathbf{P}^1}((0) + (\infty))$ admits a global section which we may call $\frac{dt}{t}$, and gluing these gives a nowhere vanishing global section of ω_{C_n}.

The smooth locus on a Néron polygon has the natural structure of a group scheme. Let C_n denote a Néron n-gon. There is a surjective map $\mathbf{P}^1 \times \mathbf{Z}/n\mathbf{Z} \to C_n$. Since $\mathbf{P}^1 - \{0, \infty\}$ is isomorphic to $\mathbf{G}_m$ and maps isomorphically to the smooth locus of each irreducible component of C_n, we get an isomorphism

$$\mathbf{G}_m \times \mathbf{Z}/n\mathbf{Z} \to C_n^{\mathrm{sm}}$$

which we use to give C_n^{sm} the structure of an S-group. The multiplication map $\mathbf{G}_m \times \mathbf{G}_m \to \mathbf{G}_m$ extends to a map $\mathbf{G}_m \times \mathbf{P}^1 \to \mathbf{P}^1$, and consequently we also have an extension

$$C_n^{\mathrm{sm}} \times C_n \to C_n$$

of the multiplication map on C_n^{sm}.

Thus, a Néron polygon with its group structure as described above is a generalized elliptic curve. When computing moduli of elliptic curves we will care about the automorphism group of such a structure; we compute

$$\mathrm{Aut}(\mathbf{G}_m \times \mathbf{Z}/n\mathbf{Z}) \cong \langle \sigma \rangle \times (\mu_n \rtimes (\mathbf{Z}/n\mathbf{Z})^\times)$$

where σ denotes the inversion on $\mathbf{G}_m \times \mathbf{Z}/n\mathbf{Z}$. Here if $\zeta \in \mu_n$ and $u \in (\mathbf{Z}/n\mathbf{Z})^\times$, the actions are

$$\sigma.(\alpha, t) = (\alpha^{-1}, -t), \qquad \zeta.(\alpha, t) = (\zeta^t \alpha, t), \qquad u.(\alpha, t) = (\alpha, ut)$$

For such an automorphism to extend to the Néron polygon, we must have $u = 1$, so we see that $\mathrm{Aut}(C_n) \cong \langle \sigma \rangle \times \mu_n$.

Proof of Proposition 1.3. Let $\pi : \widetilde{C} \to C$ be the normalization; we can write $\widetilde{C} = \bigsqcup_{i=1}^a C_i$ where each C_i is smooth and proper. Define the following quantities:

$$\begin{aligned} b &\coloneqq \#\text{ of nodes of } C \\ b_i &\coloneqq \#\text{ of points on } C_i \text{ above nodes} \\ g_i &\coloneqq \text{genus of } C_i \end{aligned}$$

We have the exact sequence

$$0 \longrightarrow \mathcal{O}_C \longrightarrow \pi_* \mathcal{O}_{\widetilde{C}} \longrightarrow \bigoplus_{j=1}^{b} k_j \longrightarrow 0$$

where k_j denotes a skyscraper for the field k at the jth node. Taking long exact sequences yields

$$0 \longrightarrow k \longrightarrow \bigoplus_{i=1}^{a} k \longrightarrow \bigoplus_{j=1}^{b} k_j \longrightarrow H^1(\mathcal{O}_C) \longrightarrow \bigoplus_{i=1}^{a} H^1(\mathcal{O}_{C_i}) \longrightarrow 0$$

so we arrive at

$$a = b + \sum_{i=1}^{a} g_i$$

and also $b \in \{a-1, a\}$. If $b = a - 1$, then $a = 1$ and $g_1 = 1$, so C is an elliptic curve.

If $b = a$, then all g_i are zero, so $C_i \cong \mathbf{P}^1$. Since ω_C is trivial, each $\omega_{C_i}(b_i)$ admits a global section, hence $b_i \geq 2$. But

$$2a = 2b = \sum_{i=1}^{a} b_i \geq 2a$$

implies $b_i = 2$ for all i. So C is a Néron a-gon. □

Morphisms in the category of generalized elliptic curves over S consist of those S-morphisms which restrict to S-group morphisms on smooth loci.

We can now define new moduli functors $\overline{F}_0, \overline{F}_1 : (\mathrm{Sch}/\mathbf{Z}[1/N])^{\mathrm{opp}} \to \mathrm{Set}$ just as before, but with elliptic curves replaced by generalized elliptic curves. To be precise, $\overline{F}_0$ parametrizes nontrivial homomorphisms whose kernel intersects every irreducible component.

It takes some additional work to show that $\overline{F}_1$ is representable (see [2]). Let $X_1(N)$ denote the resulting curve over $\mathbf{Z}[1/N]$. Again $(\mathbf{Z}/N\mathbf{Z})^\times$ acts on $X_1(N)$ and the quotient by this action yields a smooth proper curve $X_0(N)$ which serves as a coarse moduli space for $\overline{F}_0$.

Example 1.5. Let us exhibit families of elliptic curves which degenerate to Néron polygons. Let R a DVR with uniformizer π. For a 1-gon, this is easy, we can just write down

$$y^2 = x(x-1)(x-\pi)$$

which degenerates to a nodal cubic on the special fiber. Families degenerating to 2-gons and 3-gons are not much harder; we can take

$$y^2 + xy = \pi x^3 + x, \qquad y^2 + \pi y = \pi x^3 + x^2,$$

respectively. For more sides, it will be necessary to depart from plane cubics. One method of constructing such families is to begin with any Weierstrass equation that is an elliptic curve on the generic fiber, then perform blowups at the singularities on the special fiber.

On $X_1(N)$ we have the reduced relative Cartier divisor $\mathrm{Cusp}_1(N) \coloneqq X_1(N) \setminus Y_1(N)$. Similarly on $X_0(N)$ we have $\mathrm{Cusp}_0(N) \coloneqq X_0(N) \setminus Y_0(N)$. The schemes $\mathrm{Cusp}_0(N)$ and $\mathrm{Cusp}_1(N)$ are finite étale over $\operatorname{Spec} \mathbf{Z}[1/N]$, and étale locally consist of constant sections given by isogenies of standard Néron polygons [2].

It is first a bit simpler to consider the case where $N = p$ is prime. In this case, cusps on $X_1(p)$ correspond to cuspidal cubics or Néron p-gons. In the cubic case, the torsion section must be a generator of μ_p hence there are $p-1$ choices, ambiguous up to the action of σ. So there are $\frac{p-1}{2}$ cusps corresponding to cuspidal cubics. In the p-gon case, we see that $\mathrm{Aut}(C_p)$ acts transitively on the μ_p in each connected component of C_p^{sm}, so to specify a cusp is the same as specifying a connected component of C_p^{sm} up to the action of σ. Hence there are $\frac{p-1}{2}$ cusps in this case as well. In all, there are $p-1$ cusps.

When we quotient by the action of $(\mathbf{Z}/p\mathbf{Z})^\times$, each of these two sets of cusps is collapsed to a single point. So $X_0(p)$ has two cusps.

Now we say a word or two about general N. For each $d \mid N$, we have the standard Néron d-gon C_d, whose smooth locus is isomorphic to $\mathbf{G}_m \times \mathbf{Z}/d\mathbf{Z}$. We need to determine the elements P of exact order N in $\mu_N \times \mathbf{Z}/d\mathbf{Z}$ whose projection to $\mathbf{Z}/d\mathbf{Z}$ is a generator. Fix a primitive Nth root of unity ζ_0 and $a \in (\mathbf{Z}/d\mathbf{Z})^\times$. Then for $0 \le k \le N-1$ we have (ζ_0^k, a) has exact order N iff $(k, N/d) = 1$. Also (ζ_0^k, a) and $(\zeta_0^{k+mN/d}, a)$ are related by an automorphism, so it suffices to consider $0 \le k < N/d$. There are $\phi(N/d)$ choices of k in this range. After considering the automorphism σ, we see that there are $\frac{1}{2}\phi(d)\phi(N/d)$ cusps corresponding to d-gons and therefore $\mathrm{Cusp}_1(N)$ is finite of order

$$\frac{1}{2}\sum_{d \mid N} \phi(d)\phi(N/d).$$

We can conduct a similar analysis for $\mathrm{Cusp}_0(N)$. The end result will be that there are $\phi((d, N/d))$ cusps corresponding to d-gons. In particular there is a single cusp corresponding to 1-gons (labeled ∞) and one corresponding to N-gons (labeled 0).

Remark. When N is squarefree these formulas simplify; there are $\phi(N)/2$ cusps on $X_1(N)$ for each d, and on $X_0(N)$ these are collapsed to a single cusp for each d.

1.3 Correspondences

I don't have a lot to add here beyond that in [5]. For the rest of this section, I will restrict attention to the curve $X = X_0(N)$ (with open subcurve $Y = Y_0(N)$).

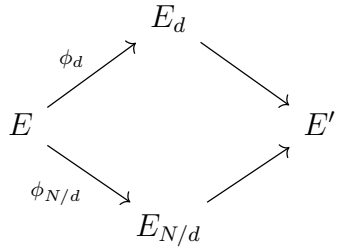
We will start with an involution on Y. We have an involution of the functor F_0 which sends an isogeny $\phi : E \to E'$ to its dual isogeny $\widehat{\phi} : E' \to E$. We correspondingly get an involution $w_N : Y \to Y$ on the coarse moduli space, and it will extend to an involution on X. As such, it must permute the cusps. Because there is an isogeny mapping the N-gon to the 1-gon (extending the projection $\mathbf{G}_m \times \mathbf{Z}/N\mathbf{Z} \to \mathbf{G}_m$), we see that the cusps 0 and ∞ are exchanged by w_N.

There are additional Atkin–Lehner involutions w_d for each $d \mid N$ such that $(d, N/d) = 1$. Given $\phi : E \to E'$ we let D and D' denote the order d subgroups of $\ker \phi$ and $\ker \widehat{\phi}$, and define

$$w_d(\phi) : E/D \to E/\ker \phi \cong E' \to E'/D'$$

Note that the kernel is $\phi^{-1}(D')/D$ which is indeed cyclic of order N (using the fact that $(d, N/d) = 1$).

An alternate description, perhaps one that is easier to work with, is as follows. We have a commuting diagram of isogenies



where the compositions along the top and bottom are both ϕ, and ϕ_d has degree d, $\phi_{N/d}$ has degree N/d. Then

$$w_d(\phi) = \phi_{N/d} \circ \widehat{\phi}_d : E_d \to E_{N/d} \tag{1.1}$$

The condition that $(d, N/d) = 1$ ensures that there is a unique diagram as above.

The other correspondences to be considered are the Hecke correspondences. Fix $m \geq 1$, then $T_m \subseteq X \times X$ is the cycle consisting of pairs

$$(\phi : E \to E', \quad \phi/C : E/C \to E'/\phi(C))$$

where C is any finite flat subgroup of order m that intersects $\ker \phi$ trivially. The two projections $\mathrm{pr}_1, \mathrm{pr}_2 : T_m \to X$ are each finite. As such, there is an endomorphism

$$T_m = (\mathrm{pr}_2)_* \, \mathrm{pr}_1^* : \mathrm{Div}(X_{\mathbf{Q}}) \to \mathrm{Div}(X_{\mathbf{Q}})$$

which restricts to an endomorphism of $J = \mathrm{Jac}(X_{\mathbf{Q}})$. We write $T_m : J \to J$ for this endomorphism as well. Let $\mathbf{T}$ denote the subalgebra of $\mathrm{End}_{\mathbf{Q}}(J)$ generated by $\{T_m\}_{m \geq 1}$.

1.4 Analytic theory

We will briefly recall the analytic description of $X(\mathbf{C})$. The points of $Y(\mathbf{C})$ parametrize isogenies $\phi : E \to E'$ of elliptic curves over $\mathbf{C}$, cyclic of degree N. Any such isogeny can be written as $\mathbf{C}/\Lambda \to \mathbf{C}/\Lambda'$, where $\Lambda \subseteq \Lambda'$ is a sublattice with $\Lambda'/\Lambda \cong \mathbf{Z}/N\mathbf{Z}$. After homothety, we can assume that $\Lambda = \langle \tau, 1 \rangle$, $\Lambda' = \langle \tau, 1/N \rangle$. For $M = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbf{Z})$ to preserve the pair (Λ, Λ') we need

$$\frac{c\tau + d}{N} \equiv \frac{a}{N} \mod \Lambda$$

for some unit a modulo N. This is equivalent to $c \equiv 0 \mod N$. Letting $\Gamma_0(N)$ denote the subgroup of $\mathrm{SL}_2(\mathbf{Z})$ with $c \equiv 0 \mod N$, we obtain the description $Y \simeq \Gamma_0(N)\backslash\mathcal{H}$. The compactification X is realized as $\Gamma_0(N)\backslash\mathcal{H}^*$, where $\mathcal{H}^* = \mathcal{H} \cup \mathbf{P}^1(\mathbf{Q})$. One must take some care in describing the complex structure at the cusps; we omit this discussion. One can give a direct calculation of the number of cusps by computing the cardinality of $\Gamma_0(N)\backslash\mathbf{P}^1(\mathbf{Q})$ (ultimately it is equivalent to the calculation with Néron polygons).

Given Λ, Λ' as above, we have a chain of inclusions $\Lambda \subseteq \Lambda' \subseteq N^{-1}\Lambda$ via

$$\langle \tau, 1 \rangle \subseteq \langle \tau, 1/N \rangle \subseteq \langle \tau/N, 1/N \rangle$$

The dual isogeny is described as $\mathbf{C}/\Lambda' \to \mathbf{C}/N^{-1}\Lambda$. After scaling by $-1/\tau$ this is described by the inclusion of lattices $\langle -1/N\tau, 1 \rangle \subseteq \langle -1/N\tau, 1/N \rangle$. Thus the operation of "dual isogeny" induces $\tau \mapsto -1/N\tau$ in the complex uniformization. We conclude that the Atkin–Lehner involution w_N acts by matrix

$$w_N = \begin{bmatrix} & -1 \\ N & \end{bmatrix} \in \mathrm{GL}_2(\mathbf{Q})^+.$$

One can similarly find matrices for the other Atkin–Lehner involutions. The involution w_d can be represented as follows. Choose $a, b \in \mathbf{Z}$ such that $ad - b(N/d) = 1$. Then

$$w_d = \begin{bmatrix} da & b \\ N & d \end{bmatrix}.$$

Hecke correspondences can also be described explicitly in the analytic picture. Consider an inclusion of lattices $(\Lambda \subseteq \Lambda')$ representing a cyclic isogeny of degree N. Applying T_m yields

$$\sum_{\substack{\Omega \cap \Lambda' = \Lambda \\ [\Omega : \Lambda] = m}} (\Omega \subseteq \Omega + \Lambda')$$

In terms of the uniformization, this translates as follows. Set

$$R_N \coloneqq \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{Mat}_2(\mathbf{Z}) : c \equiv 0 \mod N\}.$$

Then

$$T_m(\tau) = \sum_{\substack{\gamma \in \Gamma_0(N)\backslash R_N \\ \det \gamma = m}} (\gamma\tau)$$

1.5 Heegner points

Let $x = (\phi : E \to E')$ be an isogeny of elliptic curves over $\mathbf{C}$, cyclic of degree N. By abuse we identify x with the corresponding $\mathbf{C}$-valued point of Y. If $\operatorname{End} E = \operatorname{End} E' = \mathcal{O}_K$ where K is some imaginary quadratic field, we say that x is a *Heegner point (for the order $\mathcal{O} = \mathcal{O}_K$)*.

Henceforth we will fix a choice of K of discriminant D and consider only Heegner points for $\mathcal{O}$. By the theory of complex multiplication, if x is such a Heegner point, then $x \in X(H)$ where H is the Hilbert class field of K. If X has a Heegner point for $\mathcal{O}$, then $D \equiv \square \bmod 4N$ (the existence of a Heegner point implies that N is the norm of some ideal).

For simplicity, we will often make the following assumptions on K:

- $|D| > 4$ (equivalent to $K \neq \mathbf{Q}(\mu_3), \mathbf{Q}(\mu_4)$).
- D is odd.
- Every prime $p \mid N$ splits in $\mathcal{O}$.

The last assumption implies that $\mathcal{O}/\mathfrak{n} \cong \mathbf{Z}/N\mathbf{Z}$ for some ideal $\mathfrak{n}$.

Let $\mathscr{H} = \mathscr{H}(\mathcal{O}, N)$ denote the set of Heegner points on $X_0(N)(H)$. Clearly $\mathscr{H}$ admits an action of $\operatorname{Gal}(H/K) \cong \operatorname{Cl}_K$. We also see that the Atkin–Lehner involutions act on $\mathscr{H}$, and this action commutes with the Galois action.

Let $x \in \mathscr{H}$. Then x is given by a map $\mathbf{C}/\mathfrak{a} \to \mathbf{C}/\mathfrak{b}$ cyclic of degree N, where $\mathfrak{a}, \mathfrak{b}$ are fractional ideals of $\mathcal{O}$. In other words, $\mathfrak{b}/\mathfrak{a}$ should be cyclic of degree N, which is equivalent to $\mathcal{O}/\mathfrak{a}\mathfrak{b}^{-1}$ being cyclic of degree N. So to specify a Heegner point, we must give the data of an ideal class $[\mathfrak{a}] \in \operatorname{Cl}_K$ and an integral ideal $\mathfrak{n}$ with $\mathcal{O}/\mathfrak{n} \cong \mathbf{Z}/N\mathbf{Z}$. We will write $x = (\mathfrak{n}, [\mathfrak{a}])$.

Remark. Suppose d, e are relatively prime and $de = N$. Let $x_d = (\phi : E \to E')$ and $x_e = (\rho : E' \to E'')$ two isogenies of CM elliptic curves with endomorphisms by $\mathcal{O}$. The composition $x_e \circ x_d$ defines a point $x_N = (\rho \circ \phi : E \to E'')$. On pairs this is

$$(\mathfrak{e}, [\mathfrak{a}\mathfrak{d}^{-1}]) \circ (\mathfrak{d}, [\mathfrak{a}]) = (\mathfrak{e}\mathfrak{d}, [\mathfrak{a}]).$$

Proposition 1.6. *The set $\mathscr{H}$ is a torsor for $W \times \operatorname{Cl}_K$.*

Proof. The theory of complex multiplication tells us that Cl_K acts via $[\mathfrak{b}].(\mathfrak{n}, [\mathfrak{a}]) = (\mathfrak{n}, [\mathfrak{a}\mathfrak{b}^{-1}])$. So it will suffice to exhibit the W-action explicitly.

First we determine the effect on $(\mathfrak{n}, [\mathfrak{a}])$ of taking the dual isogeny. Given the isogeny $\mathbf{C}/\mathfrak{a} \to \mathbf{C}/\mathfrak{a}\mathfrak{n}^{-1}$, the dual isogeny is given by

$$\mathbf{C}/\mathfrak{a}\mathfrak{n}^{-1} \to \mathbf{C}/\mathfrak{a}N^{-1}$$

so taking dual isogenies corresponds to the operation

$$(\mathfrak{n}, [\mathfrak{a}]) \mapsto (\overline{\mathfrak{n}}, [\mathfrak{a}\mathfrak{n}^{-1}]).$$

To deal with a general Atkin–Lehner involution, write $N = p^k m$ where $(p, m) = 1$. Write $\mathfrak{n} = \mathfrak{p}^k\mathfrak{m}$ where $\mathfrak{p}$ is a prime over p. The above description of the effect of dualizing on pairs, along with the description of w_{p^k} from (1.1), yields

$$w_{p^k}(\mathfrak{n}, [\mathfrak{a}]) = (\overline{\mathfrak{p}}^k\mathfrak{m}, [\mathfrak{a}\mathfrak{p}^{-k}]).$$

□

Hecke operators also act on Heegner points, if one allows consideration of Heegner points for non-maximal orders. See [3] for details.

2 Modular curves II

In this section we will pay greater intention to the subtleties of moduli problem for elliptic curves with level structure, following Katz–Mazur [6].

2.1 Moduli problems on elliptic curves

It is desirable to have a model for $X_0(N)$ over $\mathbf{Z}$, though it is too much to hope that it will be smooth in general. Katz–Mazur offer a way to do this by considering moduli of elliptic curves with *Drinfeld level structure* rather than the usual level structure we have been using. We will state most of our results for the curves $Y_0(N)$, ignoring the cusps.

Let $\mathbf{Ell}$ denote the moduli stack of elliptic curves. For a ring A, let $\mathbf{Ell}_A$ denote the base change to A.

Definition 2.1. A *moduli problem* for $\mathbf{Ell}_A$ is a contravariant functor $\mathcal{P} : \mathbf{Ell}_A \to \mathrm{Set}$.

Definition 2.2. A moduli problem on $\mathbf{Ell}_A$ is *relatively representable* if for every elliptic curve E/A, the functor $(\mathrm{Sch}/A)^{\mathrm{opp}} \to \mathrm{Set}$ defined by $T \mapsto \mathcal{P}(E_T/T)$ is representable by an A-scheme $\mathcal{P}_{E/A}$.

In §1 we have seen examples of moduli problems, for example, $\Gamma_1(N)$ and $\Gamma_0(N)$ level structures (for $A = \mathbf{Z}[1/N]$). The notion of representability of a moduli problem is the obvious one. Given a representable moduli problem $\mathcal{P}$, we let $\mathcal{M}(\mathcal{P})$ denote its fine moduli space. Given a relatively representable moduli problem, we let $M(\mathcal{P})$ denote its coarse moduli space[1].

Definition 2.3. Given a scheme S and a an elliptic curve E/S, we define *Drinfeld level structures* as follows:

- A Drinfeld $\Gamma(N)$-structure on E is a map $\phi : (\mathbf{Z}/N\mathbf{Z})^2 \to E(S)$ such that
$$E[n] = \sum_{a,b \in \mathbf{Z}/\mathbf{N}\mathbf{Z}} [\phi(a,b)]$$
as (relative S-)Cartier divisors on E.
- A Drinfeld $\Gamma_1(N)$-structure on E is an isogeny $f : E \to E'$ and a map $\phi : \mathbf{Z}/N\mathbf{Z} \to \ker f$ such that
$$\ker f = \sum_{a \in \mathbf{Z}/N\mathbf{Z}} [\phi(a)]$$
as (relative S-)Cartier divisors on E.

[1] We can always construct $M(\mathcal{P})$ as follows. Let $\mathcal{S}$ be a representable moduli problem, finite étale Galois over $\mathbf{Ell}$ with Galois group G. Then joint $(\mathcal{P}, \mathcal{S})$-level structure is a representable moduli problem, and $M(\mathcal{P}) \cong \mathcal{M}(\mathcal{P}, \mathcal{S})/G$.

- A Drinfeld $\Gamma_0(N)$-structure on E is an isogeny $f : E \to E'$ of degree N such that $\ker f$ fppf locally on S admits a generator (as in the definition of $\Gamma_1(N)$-level structure). Equivalently, we can specify a finite flate subgroup scheme $G \subset E$ of rank N which fppf locally admits a generator.

Henceforth we will omit the adjective "Drinfeld."

Each of the above types of level structures defines a moduli problem, which we will denote $[\Gamma(N)]$, etc.

Theorem 2.4. *The moduli problems* $[\Gamma(N)]$ *and* $[\Gamma_1(N)]$ *are representable moduli problems with* $Y(N) \coloneqq \mathcal{M}(\Gamma(N))$, $Y_1(N) \coloneqq \mathcal{M}(\Gamma_1(N))$ *affine, regular, and of relative dimension* 1 *over* $\mathbf{Z}$. *The moduli problem* $[\Gamma_0(N)]$ *is relatively representable with* $Y_0(N) = M(\Gamma_0(N))$ *affine and of relative dimension* 1 *over* $\mathbf{Z}$.

Each of the three moduli problems above is finite flat over **Ell**. *Over* $\mathbf{Ell}_{\mathbf{Z}[1/N]}$ *they are finite étale.*

Our goal is to study these moduli problems over $\mathbf{F}_p$, where $p \mid N$. The main tool will be a general theorem on "crossings." The setup is as follows; we are given a field k and a commuting diagram

$$\begin{array}{ccc} \bigsqcup_{i\in I} Z_i & \longrightarrow & X \\ & \searrow & \downarrow \\ & & Y \\ & & \downarrow \\ & & k \end{array}$$

where Y is a smooth curve over k and X is finite flat over Y. For each i, the map $Z_i \to X$ is a closed immersion. We assume that

- Each Z_i is finite flat over Y.
- Each $(Z_i)^{\mathrm{red}}$ is smooth over k.

We assume that there is a given set of *supersingular points* $Y^{\mathrm{ss}} \subset Y(k)$ such that:

- For each $y \in Y^{\mathrm{ss}}$, there is a unique preimage $x \in X(k)$ such that

  $$\widehat{\mathcal{O}}_{X,x} \cong k[\![x, y]\!]/f$$

  for some $f \in k[\![x, y]\!]$.
- For each $y \in Y^{\mathrm{ss}}$ and each $i \in I$, there is a unique preimage $z_i \in Z_i(k)$.
- Over the "ordinary locus" $Y \setminus Y^{\mathrm{ss}}$ the map $\sqcup_i Z_i \to X$ is an isomorphism.

Theorem 2.5. *Let the setup be as above. Let $y \in Y^{\mathrm{ss}}$ and $x \in X(k), z_i \in Z_i(k)$ above y. Then for each i there is $f_i \in k[\![x,y]\!]$ such that $\widehat{\mathcal{O}}_{Z_i,z_i} \cong k[\![x,y]\!]/f_i^{e_i}$, and furthermore we have*

$$\widehat{\mathcal{O}}_{X,x} \cong k[\![x,y]\!]/\prod_i f_i^{e_i}.$$

If Y is connected, then so is each Z_i, and then the Z_i's are the irreducible components of X.

We will not give a proof here, but ignoring the statements about complete local rings, we can at least justify the last assertion (that if Y is connected then so is each Z_i and the Z_i's are the components of X). This is the easy part of the theorem, but is also the statement that will be most important to us.

To justify that the Z_i are connected, observe that each Z_i is finite flat over Y and that there is a unique preimage over each supersingular point. This immediately precludes the possibility of multiple connected components. Once Z_i is connected, it is irreducible since we have assumed its reduced subscheme is smooth. Now the Z_i are the irreducible components of X since the map $\sqcup_i Z_i \to X$ is an isomorphism over $Y \setminus Y^{\mathrm{ss}}$.

2.2 Prime level

For the rest of this section we will work entirely over $\mathbf{Ell}_{\mathbf{F}_p}$ (so e.g. $[\Gamma_0(p^n)]$ refers to a moduli problem over $\mathbf{F}_p$). We will start by examining the reduction $[\Gamma_0(p)]$ modulo p. Though the theory for prime power level will still be tractable, it is still instructive to first look at this special case where things are simpler.

By definition now $[\Gamma_0(p)]$ is the moduli stack of p-isogenies $\phi : E \to E'$ over base schemes $S/\mathbf{F}_p$. Given an elliptic curve E/S, we let $F : E \to E^{(p)}$ denote the relative Frobenius over S, and $V : E^{(p)} \to E$ its dual isogeny.

Lemma 2.6. *Any p-isogeny of ordinary elliptic curves is of the form $F : E \to E^{(p)}$ or $V : E^{(p)} \to E$, up to isomorphism.*

Proof. First suppose E is ordinary and $\phi : E \to E'$ is a p-isogeny. Then $\ker F$ is étale locally isomorphic to μ_p. Let $H = \ker \phi$. Then $H = \ker F$ or $H \cap \ker F = \{O\}$. If $H = \ker F$, we're done. If $H \cap \ker F = \{O\}$, the composite map

$$H \longrightarrow E[p] \xrightarrow{F} \ker(V : E^{(p)} \to E)$$

is an isomorphism. So H is étale locally isomorphic to $\mathbf{Z}/p\mathbf{Z}$. Let $H' = \ker(\widehat{\phi} : E' \to E)$. Then H' is étale locally μ_p, so $H' \cong \ker(F : E' \to E'^{(p)})$. We conclude that $E \cong E'^{(p)}$ and ϕ is up to automorphisms $V : E'^{(p)} \to E'$. □

Lemma 2.7. *Let $\iota_F, \iota_V : \mathbf{Ell}_{\mathbf{F}_p} \to [\Gamma_0(p)]$ be defined via*

$$E/S \mapsto (F : E \to E^{(p)}), (V : E^{(p)} \to E),$$

respectively. Then ι_F, ι_V are closed immersions.

The composite maps

$$\mathbf{Ell}_{\mathbf{F}_p} \xrightarrow{\iota_F} [\Gamma_0(p)] \longrightarrow \mathbf{Ell}_{\mathbf{F}_p}$$
$$\mathbf{Ell}_{\mathbf{F}_p} \xrightarrow{\iota_V} [\Gamma_0(p)] \longrightarrow \mathbf{Ell}_{\mathbf{F}_p}$$

have degrees 1 and p, respectively.

Lemma 2.8. *Let k a field of characteristic p. Then any p-isogeny of supersingular elliptic curves is the Frobenius morphism (up to isomorphism).*

Proof. In the supersingular case, any p-isogeny is inseparable, hence must factor through the Frobenius. □

Theorem 2.9. *Let $\mathcal{S}$ be a representable moduli problem. Then $\iota_F, \iota_V : \mathcal{M}(\mathcal{S}) \to \mathcal{M}(\mathcal{S}, \Gamma_0(p))$ as above are closed immersions, and $\mathcal{M}(\mathcal{S}, \Gamma_0(p))$ is the disjoint union of the two copies of $\mathcal{M}(\mathcal{S})$ with crossings at the supersingular points.*

Proof. Follows from the preceding lemmas and the crossings theorem. □

Corollary 2.10. *Let $\iota_F, \iota_V : \mathbf{A}^1_j \to Y_0(p)$ the induced maps on coarse moduli spaces. Then ι_F, ι_V are closed immersions and $Y_0(p)$ is the disjoint union of the two copies of $\mathbf{A}^1_j$ with crossings at the supersingular points.*

More generally, Theorem 2.9 shows that for any m coprime to p, we have that $Y_0(mp)$ is the disjoint union of two copies of $Y_0(m)$ with crossings at supersingular points.

Corollary 2.11. *The scheme $Y_0(p)/\mathbf{Z}$ fails to be regular only at crossings on the fiber at p with automorphism group larger than $\{\pm 1\}$.*

For example, $Y_0(13)/\mathbf{Z}$ is regular.

2.3 Prime power level

By looking at $[\Gamma_0(p)]$ modulo p we have done enough to handle $[\Gamma_0(N)]$ over $\mathbf{Z}$ when N is squarefree. We will later feel free to assume that N is squarefree, so this is all we need. However, we will anyway briefly address the case that a prime power p^n divides N (without proofs, which can be found in [6]). So we need to examine $[\Gamma_0(p^n)]$ modulo p. Again, from now on we'll consider all moduli problems modulo p.

As an intermediate step, it is useful to consider the moduli problem $[p^n-\mathrm{Isog}]$ of p^n-isogenies (not necessarily cyclic). Similarly to the case of $n = 1$, we can describe such isogenies very explicitly.

Lemma 2.12. *Any p^n-isogeny of ordinary elliptic curves factors uniquely as*

$$E \xrightarrow{F^a} E^{(p^a)} \cong E'^{(p^b)} \xrightarrow{V^b} E'$$

for a unique choice of $a, b \geq 0$ with $a + b = n$.

Let $[(a,b)]$ denote the moduli problem of p^n-isogenies that factor as in Lemma 2.12 (henceforth called (a,b)-*isogenies*, note that we allow supersingular curves though the lemma is stated for ordinary curves). For any $d \geq 0$, let $\iota_d : \mathbf{Ell}_{\mathbf{F}_p} \to \mathbf{Ell}_{\mathbf{F}_p}$ denote the functor $E \mapsto E^{(p^d)}$. From Lemma 2.12 we see that $[(a,b)]$ can be described as $(\iota_a \times \iota_b)^{-1}(\Delta)$ where Δ is the diagonal in $\mathbf{Ell}_{\mathbf{F}_p} \times \mathbf{Ell}_{\mathbf{F}_p}$. If $\mathcal{S}$ is a representable moduli problem on $\mathbf{Ell}_{\mathbf{F}_p}$, this concretely means that $\mathcal{M}(\mathcal{S}, [(a,b)]) \cong (\iota_a \times \iota_b)^{-1}(\Delta)$ where Δ is the diagonal in $\mathcal{M}(\mathcal{S}) \times \mathcal{M}(\mathcal{S})$.

Proposition 2.13. *Let $a, b \geq 1$. An (a,b) isogeny $E \to E'$ of ordinary elliptic curves over S is cyclic iff there is a closed subscheme $S_0 \hookrightarrow S$ so that S is a thickening of S_0 of order $\leq p-1$, such that $E_{S_0}^{(p^{a-1})} \cong E_{S_0}^{(p^{b-1})}$ compatibly with the isomorphism $E^{(p^a)} \cong E^{(p^b)}$.*

For a proof see [6, Theorem 13.3.5]. I do not have something both short and convincing to say, but I will anyway say a little for $a = b = 1$. In this case the diagram is

$$E \xrightarrow{F} E^{(p)} \xrightarrow{\varepsilon} E'^{(p)} \xrightarrow{V} E'$$

where ε is an isomorphism. The defining equation for the coarse moduli space of p^2-isogenies as a subscheme of $\mathbf{A}^1_j \times \mathbf{A}^1_{j'}$ should be $j^p = j'^p$; however if $j = j'$ the given isogeny is $[p]$ which is not cyclic. The defining equation for $Y_0(p^2; 1, 1)$ ends up being $(j - j')^{p-1}$.

For $a, b \geq 1$, let $[\Gamma_0(p^n); a, b]$ denote the moduli problem determined by $(\iota_a \times \iota_b)^{-1}(\mathrm{Inf}^{p-1}(\Delta))$. If $a = 0$ or $b = 0$, instead take $(\iota_a \times \iota_b)^{-1}(\Delta)$.

Theorem 2.14. *The natural map $[(a,b)] \to [p^n\text{−Isog}]$ is a closed immersion. The diagram*

$$\begin{array}{ccc} [\Gamma_0(p^n); a, b] & \longrightarrow & [\Gamma_0(p^n)] \\ \downarrow & & \downarrow \\ [(a,b)] & \longrightarrow & [p^n\text{−Isog}] \end{array}$$

is Cartesian, and in particular the top arrow is a closed immersion as well.

Corollary 2.15. *Let $\mathcal{S}$ be a representable moduli problem. Then $\mathcal{M}(\mathcal{S}, [\Gamma_0(p^n)])$ is the disjoint union of $\mathcal{M}(\mathcal{S}, [\Gamma_0(p^n); a, b])$ with crossings at the supersingular points.*

We then get similar descriptions of the coarse moduli spaces, as we did for prime level.

It remains to ascertain the geometric properties of the natural map $[\Gamma_0(p^n); a, b] \to \mathbf{Ell}_{\mathbf{F}_p}$ (the map remembering the source E of the isogeny). Suppose $a, b \geq 1$. This map is still finite flat. Let us again start by considering $a = b = 1$. On the level of coarse moduli spaces over $\mathbf{F}_p$, we have seen that this corresponds to the inclusion of rings

$$\mathbf{F}_p[x] \hookrightarrow \mathbf{F}_p[x, y]/(x - y)^{p-1}.$$

In particular we will see that $Y_0(p^2; 1, 1)$ has degree $p - 1$ over $\mathbf{A}^1_j$, and that $Y_0(p^2; 1; 1)$ is nonreduced of multiplicity $p - 1$, with $Y_0(p^2; 1; 1)^{\mathrm{red}} \cong \mathbf{A}^1_j$.

Similarly, for $[\Gamma_0(p^n); a, b]$ with $a, b \geq 1$, the corresponding inclusion is

$$\mathbf{F}_p[x] \hookrightarrow \mathbf{F}_p[x, y]/(x^{p^{a-1}} - y^{p^{b-1}})^{p-1}$$

which has degree $p^{b-1}(p-1)$ and is nonreduced of multiplicity $p^{\min\{a,b\}-1}(p-1)$. If $a \geq b$ the underlying reduced subscheme is cut out by $x^{p^{a-b}} - y$; similarly for $a \leq b$.

For $a = 0$ we get defining equation $x - y^{p^n}$ so the coarse scheme is reduced (itself isomorphic to $\mathbf{P}^1_j$) of degree p^n over $\mathbf{P}^1_j$. For $b = 0$ we get $x^{p^n} - y$ so again the coarse scheme is $\mathbf{P}^1_j$ and has degree 1 over $\mathbf{P}^1_j$.

2.4 Cusps

Let us again work over $\mathbf{Z}$ and restrict to the case of $X_0(p)$. We omit the discussion of how to define the moduli problem over $\mathbf{Z}$, it is completely analogous to how it was done in §1 except now we use Drinfeld level structure. There are still two cusps which are given by sections $\operatorname{Spec}\mathbf{Z} \to X_0(p)$ corresponding to Néron polygons; as before ∞ denotes the p-gon cusp and 0 denotes the 1-gon cusp.

In the case of $X_0(p)$ we have seen that on the fiber at p there are only two irreducible components $X_0(p;1,0)$ and $X_0(p;0,1)$, each isomorphic to $\mathbf{P}^1_j$. We have seen that the p-gon cusp has kernel $\mathbf{Z}/p\mathbf{Z}$ while the 1-gon cusp has kernel μ_p; it is then clear that on the fiber at p we have $\infty \in X_0(p;1,0)$ and $0 \in X_0(p;0,1)$.

Similar statements can be made for general $X_0(N)$; see [5, §III.1].

3 L-functions of modular forms

3.1 L-functions and modularity

In the historical development of this subject, people including (but not limited to) Birch, Stephens, and Swinnerton–Dyer had noticed relationships between the structure of the group $E(\mathbf{Q})$ and "central values" of the L-function $L(E,s)$ for an elliptic curve $E/\mathbf{Q}$. Of course, for the term "central value" to make sense, $L(E,s)$ must possess a functional equation, which was only known in the case that E is *modular*. In this section we will explain the meaning of this term. (Of course following the work of Wiles and subsequent developements, it is known that all $E/\mathbf{Q}$ are modular). In this section, all modular forms for $\Gamma_0(N)$ will be of weight 2 unless otherwise specified.

Definition 3.1. An elliptic curve $E/\mathbf{Q}$ is *modular* if $L(E,s) = L(f,s)$ for some newform f for $\Gamma_0(N)$.

Definition 3.2. An elliptic curve $E/\mathbf{Q}$ is *modular* if there is a nonconstant morphism $X_0(N) \to E$ for some N. Equivalently, there is a nonconstant homomorphism $J_0(N) \to E$ for some N.

We usually normalize the map $X_0(N) \to E$ so that $\infty \mapsto O$, and consider the Abel–Jacobi map $X_0(N) \hookrightarrow J_0(N)$ with respect to the point ∞.

These two definitions are equivalent, but this is not trivial. Let us at least briefly discuss how one obtains a modular form from a homomorphism $J_0(N) \to E$.

We will now fix $N \geq 1$ and return to the notation $X = X_0(N)$.

Let $\mathfrak{N}_0(N)$ denote the set of normalized cuspidal Hecke eigenforms for $\Gamma_0(N)$. For $f \in \mathfrak{N}_0(N)$ we let $[f]$ denote its class in $\mathfrak{N}_0(N)/G_{\mathbf{Q}}$.

Recall that $\mathbf{T} \subseteq \operatorname{End}_{\mathbf{Q}}(J_0(N))$ is the algebra of Hecke operators. For each $f \in \mathfrak{N}_0(N)$, we obtain a character $\lambda_f : \mathbf{T} \to \mathbf{C}$. Set $M_f \mid N$ so that f is a newform for $\Gamma_0(M_f)$. We set $I_f = \ker \lambda_f$ and consider the abelian variety

$$A_f \coloneqq J_0(M_f)/I_f J_0(M_f)$$

By construction, there is an action by $R_f \coloneqq \mathbf{T}/I_f \cong \mathbf{Z}[\{a_n(f)\}]$ which is finite over $\mathbf{Z}$.

Theorem 3.3. *The Jacobian $J_0(N)$ is isogenous to the product*

$$\bigoplus_{[f]\in\mathfrak{N}_0(N)/G_{\mathbf{Q}}} A_f^{\sigma_0(N/M_f)}.$$

This is essentially a consequence of the decomposition of $S_2(\Gamma_0(N))$ as a direct sum of spaces spanned by eigenforms $f(mz)$ where $m \mid (N/M_f)$.

Corollary 3.4. *An elliptic curve admits a modular parametrization iff there is some N and some newform f for $\Gamma_0(N)$ for which there is a nonconstant map $A_f \to E$.*

The equivalence between the two definitions is now easy to state; given a nonconstant map $X_0(N) \to E$ we take f so that $A_f \to E$ is nonconstant. To prove that f with the same L-function as E yields a nonconstant map $A_f \to E$ takes some work, which we will omit. Even further work is needed for the following refinement.

Theorem 3.5. *Suppose f is a newform for $\Gamma_0(N)$ and $A_f \twoheadrightarrow E$. Then $N_E = N$.*

3.2 Quadratic extensions and Hecke characters

Let K an imaginary quadratic field, and $\chi : G_K \to \mathbf{C}^\times$ a nontrivial unramified character. By class field theory, we get an unramified automorphic character $\chi : K^\times\backslash\mathbf{A}_K^\times \to \mathbf{C}^\times$, or equivalently a character of the class group Cl_K. Now $\mathrm{Ind}_K^{\mathbf{Q}}\chi$ is a 2-dimensional representation of $\mathbf{G}_{\mathbf{Q}}$, and Langlands tells us that correspondingly we should obtain an automorphic representation π_χ of GL_2 over $\mathbf{Q}$.

Lemma 3.6. *When $\chi^2 \neq 1$ the representation π_χ is cuspidal.*

The corresponding fact on the Galois side is that when $\chi^2 \neq 1$, the representation $\mathrm{Ind}_K^{\mathbf{Q}}\chi$ is irreducible. This fact is equivalent to showing that if τ is complex conjugation, then the twisted character χ^τ of G_K given by $\chi^\tau(g) = \chi(\tau g \tau^{-1})$ is not the same as χ. In fact by class field theory we have $\chi^\tau = \overline{\chi}$, so we get what we want.

Gross tells us how to write down a Hecke eigenform g_χ which will generate the representation π_χ. For ease I will assume that $K \neq \mathbf{Q}(\mu_3), \mathbf{Q}(\mu_4)$.

For each ideal class $\mathscr{A} \in \mathrm{Cl}_K$ we consider the θ-series

$$\theta_{\mathscr{A}} = \frac{1}{2} + \sum_{\substack{\mathfrak{a}\in\mathscr{A}\\ \mathfrak{a}\text{ integral}}} q^{N(\mathfrak{a})} = \sum_{n\geq 0} r_{\mathscr{A}}(n) q^n$$

where for $n \geq 1$, $r_{\mathscr{A}}(n)$ denotes the number of integral ideals in $\mathscr{A}$ of norm n, and $r_{\mathscr{A}}(0) = \frac{1}{2}$. Let $\mathfrak{b}$ be an integral ideal in $-\mathscr{A}$. The integral ideals in the class $\mathscr{A}$ are parametrized by

$$\{x \in F^\times/\{\pm 1\} : x\mathfrak{b}^{-1} \subseteq \mathcal{O}\}$$

so we get

$$\theta_{\mathscr{A}} = \frac{1}{2} + \frac{1}{2}\left(\sum_{x\in\mathfrak{b}-\{0\}} q^{N(x\mathfrak{b}^{-1})}\right) = \frac{1}{2}\sum_{x\in\mathfrak{b}} q^{N(x)/N(\mathfrak{b})}$$

Now $\mathfrak{b}$ is a rank 2 lattice equipped with the norm form, so this is the usual θ-function of a rank 2 lattice. By the general theory for lattices, it is not hard to check that it yields a modular form of weight 1 for $\Gamma_1(D)$ with quadratic character α_D corresponding to the extension $K/\mathbf{Q}$.

We now consider the series

$$g_\chi = \sum_{\mathfrak{a} \text{ integral}} \chi(\mathfrak{a}) q^{N(\mathfrak{a})} = \sum_{\mathscr{A} \in \mathrm{Cl}_K} \chi(\mathscr{A}) \theta_{\mathscr{A}}$$

From this it follows that g_χ is a modular form of weight 1 for $\Gamma_1(D)$ with character α_D.

Now let f be a newform for $\Gamma_0(N)$. We define

$$L_{\mathscr{A}}(f, s) \coloneqq L^{(N)}(\alpha_D, 2s-1) \cdot \sum_{n \geq 0} a_n r_{\mathscr{A}}(n) n^{-s}$$
$$L(f, \chi, s) \coloneqq \sum_{\chi} \chi(\mathscr{A}) L_{\mathscr{A}}(f, s),$$

where $L^{(N)}$ denotes the product of Euler factors over places prime to N. We define

$$L^*_{\mathscr{A}}(f, s) = (2\pi)^{-2s} (N|D|)^s \Gamma(s)^2 L_{\mathscr{A}}(f, s)$$

and likewise define $L^*(f, \chi, s)$. Observe that the above L-functions are closely related to the Rankin–Selberg convolutions $L(f \otimes \theta_{\mathscr{A}}, s)$ and $L(f \otimes g_\chi, s)$. This will be the main tool for proving the relevant functional equations, etc.

By some standard arguments (which we will expound later), it can be seen that $L^*_{\mathscr{A}}$ satisfies the functional equation

$$L^*_{\mathscr{A}}(f, s) = -\alpha_D(N) L^*_{\mathscr{A}}(f, 2-s)$$

and hence so does $L^*(f, \chi, 1)$. In our preferred situation when $D \equiv \square \mod 4N$, this implies that $L(f, \chi, 1) = 0$.

3.3 The Gross–Zagier formula

Fix $N \geq 1$ and let $X = X_0(N)$, $J = J_0(N)$. Let $K/\mathbf{Q}$ an imaginary quadratic field with assumptions as in §1.5. Let $x \in X(H)$ be a Heegner point, and let $c = \mathrm{AJ}(x) \in J(H)$ its image under the Abel–Jacobi map.

Since J is isogenous to a product $\bigoplus A_f^{m_f}$, we see that

$$J(H) \otimes \mathbf{C} \cong \bigoplus_{[f]} V_f^{m_f} \cong \bigoplus_{[f], \chi} V_{f,\chi}^{m_f}$$

as complex representations of $\mathrm{Gal}(H/K) \cong \mathrm{Cl}_K$, where $V_f = A_f(H) \otimes \mathbf{C}$ and χ runs over ideal class characters.

Let $c_{f,\chi} = h_K \cdot \mathrm{pr}_{V_{f,\chi}}(c)$ (the constant factor is just a convention).

Theorem 3.7 (Gross–Zagier)**.** *Let f be a newform for $\Gamma_0(N)$ and $x \in X(H)$ a Heegner point for $\mathcal{O}_K$. With setup as above,*

$$L'(f, \chi, 1) = \frac{8\pi^2 (f, f)}{h_K \sqrt{D}} \hat{h}_{2\Theta}(c_{\chi, f})$$

where $\hat{h}_{2\Theta}$ denotes the Néron–Tate canonical height on J/H associated to the divisor 2Θ.

Remark. The divisor 2Θ is characterized as follows: let Θ_0 denote the divisor on $\mathrm{Pic}^{g-1}(C)$ given as the image of the map $\mathrm{Sym}^{g-1} C \to \mathrm{Pic}^{g-1}(C)$. Then $2\Theta = 2\Theta_0 - \omega_C$ is a divisor on the Jacobian. It is both ample and symmetric, so $\hat{h}$ is quadratic.

As of now, we have been given no reason to suspect any relationship between the L-function on the LHS and the Heegner point on the RHS. The entire purpose of these notes is to explicate this relationship, but for now we just offer one piece of evidence that they are related.

Proposition 3.8. *For $(m, N) = 1$, the multiplicity of $\mathscr{A}.x$ in $T_m x$ is $r_{\mathscr{A}}(m)$.*

Proof. Suppose we are given an elliptic curve E with CM by $\mathcal{O}_K$, so $E \cong \mathbf{C}/\mathfrak{b}$ for some integral ideal $\mathfrak{b}$ of $\mathcal{O}_K$. If $\mathfrak{a}$ is a fractional ideal representative of $\mathscr{A}^{-1}$, then $\mathscr{A}.E \cong \mathbf{C}/\mathfrak{b}\mathfrak{a}$. To exhibit an isogeny $E \to \mathscr{A}.E$ of degree m is then the same as to exhibit an inclusion $\alpha^{-1}\mathfrak{b} \subseteq \mathfrak{b}\mathfrak{a}$ of index m for some $\alpha \in K$, or equivalently an $\alpha \in K$ for which $(\alpha\mathfrak{a})^{-1}$ is integral of norm m. So the desired multiplicity is $r_{-\mathscr{A}}(m) = r_{\mathscr{A}}(m)$. □

Thus the series $\theta_{\mathscr{A}}$ can be expressed in terms of Heegner points and Hecke operators.

We can reduce the Gross–Zagier theorem to a version involving the L-function $L_{\mathscr{A}}(f, s)$. Let σ denote the element of $\mathrm{Gal}(H/K)$ corresponding to $\mathscr{A}$.

Theorem 3.9 (Gross–Zagier for fixed ideal class)**.** *The q-series*

$$g_{\mathscr{A}} \coloneqq \sum_{m \geq 1} \langle c, T_m c^{\sigma} \rangle q^m$$

is a cusp form of weight 2 for $\Gamma_0(N)$. With notation as before,

$$L'_{\mathscr{A}}(f, 1) = \frac{8\pi^2}{\sqrt{D}} (f, g_{\mathscr{A}}).$$

The assertion that $g_{\mathscr{A}}$ is a cusp form follows from the following more general result:

Proposition 3.10. *Let $\alpha : \mathbf{T} \to \mathbf{C}$ any $\mathbf{Q}$-linear character. Then $\sum_{m \geq 1} \alpha(T_m) q^m$ is a cusp form of weight 2 on $\Gamma_0(N)$.*

Proof. Let $J = J_0(N)_{\mathbf{Q}}$ and recall the algebra of Hecke operators $\mathbf{T} \subseteq \mathrm{End}_{\mathbf{Q}}(J)$. We have a canonical identification $\mathrm{Lie}(J) \simeq S_2(\Gamma_0(N), \mathbf{Q})$ from which it is evident that the natural map

$$\mathbf{T} \to \mathrm{End}_{\mathbf{Q}}(\mathrm{Lie}(J))$$

is injective. We define a pairing

$$\mathbf{T} \times \mathrm{Lie}(J) \to \mathbf{Q}$$

via $(T, f) \mapsto a_1(Tf)$. We claim that this is a perfect pairing. Indeed for fixed T, suppose $a_1(Tf) = 0$ for all f. Applying this to $T_m f$, we see that $a_m(Tf) = 0$ for all f and all m, so $Tf = 0$ for all f, so $T = 0$ (from the aforementioned injectivity of $\mathbf{T} \to \mathrm{End}_{\mathbf{Q}}(\mathrm{Lie}(J))$). For fixed f, if $a_1(Tf) = 0$ for all T, then $a_m(f) = 0$ for all m, so $f = 0$.

Since this pairing is perfect, there is some $f \in S_2(\Gamma_0(N), \mathbf{C})$ for which $\alpha(T) = a_1(Tf)$, and then f is the desired cusp form. □

We will later see that Theorem 3.9 implies Theorem 3.7.

3.4 Implications for BSD

Recall the simplest statement of the BSD conjecture.

Conjecture 3.11. *Let $E/\mathbf{Q}$ be an elliptic curve. Then*

$$\mathrm{ord}_{s=1} L(E, s) = \mathrm{rank}_{\mathbf{Z}} E(\mathbf{Q})$$

i.e. the analytic rank $r_{\mathrm{an}}(E)$ of E coincides with the Mordell–Weil rank $r_{\mathrm{MW}}(E)$ of E.

There are further refinements of this conjecture (i.e. explicit conjectural formulae for $L^{(r)}(E, s)$ in terms of the Mordell–Weil lattice).

The Gross–Zagier formula gives some modest but still notable progress towards this.

Theorem 3.12. *If $r_{\mathrm{an}}(E) = 1$ then $r_{\mathrm{MW}}(E) \geq 1$.*

Proof. Let f be a normalized newform of level $N = N_E$ so that $L(E, s) = L(f, s)$; then there is a modular parametrization $\pi : X_0(N) \to E$ such that $\pi^* \omega = c\omega_f$ for some $c \in \mathbf{Q}$.

Let $x \in X_0(N)(H)$ be a Heegner point on $X_0(N)$, let $Q = \pi(x) \in E(H)$, and let

$$P = \mathrm{Tr}_{H/K} Q \coloneqq \sum_{\sigma \in \mathrm{Gal}(H/K)} Q^\sigma = \pi(c_{1,f})$$

so $P \in E(K)$. Then

$$\hat{h}_E(P) = \hat{h}(c_{1,f}) \cdot \deg \pi$$

Now we use the Gross–Zagier formula to conclude:

$$L'(f, \chi_{\mathrm{triv}}, 1) = C_0 \hat{h}_E(P)$$

where C_0 is a nonzero constant.

On the other hand,

$$L(f, \chi_{\mathrm{triv}}, s) = L(f, s) L(f \otimes \alpha_D, s) = L(E, s) L(E^D, s)$$

where α_D is the quadratic character associated to K. Since $r_{\mathrm{an}}(E) = 1$ we see that

$$L'(E, 1) L(E^D, 1) = C_0 \hat{h}_E(P)$$

According to a theorem of Waldspurger (apparently) we can choose K so that $L(E^D, 1) \neq 0$. In that situation, we see that $\hat{h}_E(P) \neq 0$ so $\mathrm{Tr}_{K/\mathbf{Q}} P$ is a point of $E(\mathbf{Q})$ of infinite order. □

4 Local heights on curves

In this section we will give some basic definitions of arithmetic intersection theory, so as to be able to understand nonarchimedean local heights on curves. The substantial proofs are omitted, we refer the reader to [7] and [8] (all the material on arithmetic surfaces comes from here). All the material on local heights comes from [4].

4.1 Arithmetic surfaces

Let $(R, \mathfrak{m} = (\pi), k)$ a DVR with fraction field F.

Definition 4.1. An *arithmetic surface* is a scheme $\mathcal{X}$ that is integral, normal, excellent, is flat and finite type over R, and such that the generic fiber $\mathcal{X}_F$ is a smooth curve over F.

Lang [7] replaces the assumption of "excellent" with the assumption that any normalization of $\mathcal{X}$ in a finite extension of its function field is finite over $\mathcal{X}$. In any case, this assumption is largely a technical one that we can ignore in all applications.

Proposition 4.2. *Let $\mathcal{X}/R$ a regular arithmetic surface. Then every point in the image of the natural map*

$$\mathcal{X}(R) \to \mathcal{X}_{\mathfrak{m}}(k)$$

is a smooth point of $\mathcal{X}_{\mathfrak{m}}/k$.

Corollary 4.3. *Let $\mathcal{X}/R$ a regular arithmetic surface and $\mathcal{X}^{\mathrm{sm}}$ the smooth locus over R. Then the natural map $\mathcal{X}^{\mathrm{sm}}(R) \to \mathcal{X}(R)$ is a bijection.*

Definition 4.4. Let X/F a smooth curve. A *model* for X is an arithmetic surface $\mathcal{X}/R$ and an isomorphism $X \simeq \mathcal{X}_F$.

Theorem 4.5. *A smooth proper curve X/F admits a proper regular model $\mathcal{X}/R$.*

A smooth model need not exist.

Henceforth in this section assume that $\mathcal{X}/R$ is a proper regular arithmetic surface, with generic fiber $X = \mathcal{X}_F$.

Definition 4.6. A *horizontal divisor* on $\mathcal{X}$ is a Weil divisor that arises as the Zariski closure of a Weil divisor on X.

A *vertical divisor* or *fibral divisor* on $\mathcal{X}$ is a Weil divisor of $\mathcal{X}$ supported on the special fiber.

Note that any irreducible Weil divisor on $\mathcal{X}$ is either horizontal or vertical.

4.2 Intersections on arithmetic surfaces

For this entire subsection, we continue to assume that $\mathcal{X}/R$ is a proper regular arithmetic surface. In particular, recall that all Weil divisors on $\mathcal{X}$ are combinations of horizontal and vertical ones.

Definition 4.7. Let D, E two effective divisors on $\mathcal{X}$ without common component, and let $x \in \mathcal{X}$ a closed point contained in both. We define the *intersection multiplicity at* x

$$i_x(D, E) \coloneqq \text{length}_{R_x}(R_x/(f, g))$$

where f, g are equations locally cutting out D, E in a neighborhood of x. We define their *intersection*

$$D \cdot E = \sum_x i_x(D, E)[x]$$

and use the shorthand

$$(D \cdot E) = \deg D \cdot E = \sum_x i_x(D, E)$$

for the total degree of the intersection.

The pairing i_x is bilinear (check this for yourself) so all the quantities defined above extend to any pair of divisors with disjoint supports.

However, these pairings do not respect linear equivalence. If R is a DVR we can consider in $\mathbf{P}^1_R = \operatorname{Proj} R[X, Y]$ the hyperplanes $\Gamma_1, \Gamma_2, \Gamma_3$ cut out by $X, X + \pi Y, Y$, respectively. Note that Γ_1 and Γ_2 intersect on the special fiber at $[0 : 1]$ with multiplicity 1, but Γ_1 and Γ_3 do not intersect.

Let $\text{Div}_f(\mathcal{X})$ denote the group of fibral divisors. It is clear that a principal divisor will pair to 0 with a fibral divisor. More precisely, if D is principal and E is fibral (say reduced and irreducible), then $D \cdot E$ is a principal divisor on E. So we have the following.

Theorem 4.8. *There is a (unique) bilinear pairing* $\text{Div}(\mathcal{X}) \times \text{Div}_f(\mathcal{X}) \to \mathbf{Z}$ *which extends the above pairing* $(D, E) \mapsto (D \cdot E)$ *and which respects linear equivalence in the first slot.*

Perhaps more precisely, we have exhibited an intersection form

$$\text{CH}_1(\mathcal{X}) \times \text{Div}_f(\mathcal{X}) \to \text{CH}_0(\mathcal{X})$$

and we can compose with the degree map $\text{CH}_0(X) \to \mathbf{Z}$.

The only nontrivial thing to compute in the above pairing is the self-intersection of a fibral component. Write $\mathcal{X}_\mathfrak{m} = \sum_i n_i F_i$ where each F_i is a proper irreducible curve over k. Now $\mathcal{X}_\mathfrak{m}$ is a principal divisor, so it follows that $(\mathcal{X}_\mathfrak{m} \cdot F_i) = 0$. This allows us to compute the self-intersection (F_i^2) in terms of the other intersection numbers $(F_i \cdot F_j)$; in particular $(F_i^2) \leq 0$.

Let $G_i = n_i F_i$ denote an irreducible component of $\mathcal{X}_\mathfrak{m}$. Let A denote the matrix with $A_{ij} = (G_i \cdot G_j)$. Observe that all rows and all columns of A sum to zero, and entries are nonnegative off the diagonal.

Proposition 4.9. *A is negative semidefinite (over $\mathbf{R}$) with kernel spanned by $(1, 1, \cdots, 1)$.*

Proof. For any $x = (x_i)$ we have

$$x^\top A x = \sum_{i,j} A_{ij} x_i x_j = -\frac{1}{2} \sum_{i,j} A_{ij} (x_i - x_j)^2 \leq 0.$$

Suppose $x_i \neq x_j$ for some $i \neq j$. Choose a sequence of indices $i = i_0, i_1, \ldots, i_k = j$ such that G_{i_m} intersects $G_{i_{m+1}}$. Consider the sum

$$\sum_{m=1}^{k} A_{i_{m-1} i_m} (x_{i_m} - x_{i_{m-1}})^2$$

Each term is nonnegative and at least one is strictly positive, so $x^\top A x < 0$. □

Corollary 4.10. *The intersection pairing on* $\mathbf{Q}\operatorname{Div}_f(\mathcal{X})/\mathbf{Q}\mathcal{X}_m$ *is negative definite.*

4.3 Construction of local heights

Let F_v be a local field with absolute value $|\cdot|_v$ and corresponding valuation $v(x) = -\log_{q_v} |x|_v$. Let X/F_v a smooth projective variety. Note that the set $X(F_v)$ comes equipped with a canonical topology induced by the topology on F_v.

Definition 4.11. A *local height pairing* for X/F_v is a real-valued symmetric bilinear map which takes as input two relatively prime divisors $D, E \in \operatorname{Div}^0(X)(F_v)$ (Weil divisors over F_v) and outputs the number $\langle D, E \rangle_v$, with the following properties:

1. For $D = (f)$ principal and $E \in Z^0(X(F_v))$, $\langle D, E \rangle_v = \log |f(E)|_v$.
2. Symmetry: $\langle D, E \rangle_v = \langle E, D \rangle_v$.
3. For $D \in \operatorname{Div}^0(X)(F_v)$ and $P_0 \notin \operatorname{supp} D$, the map $\lambda_{D,P_0} : X(F_v) - \operatorname{supp} D \to \mathbf{R}$ given by
$$P \mapsto \langle D, [P] - [P_0] \rangle_v$$
is continuous.

(Note that the last condition does not depend on the choice of P_0, for given different P_0' the two maps differ by a constant).

Theorem 4.12 (Néron)**.** *Local heights for curves exist and are unique.*

Let L_w/F_v be a finite extension of degree $r = ef$ (ramification and inertia). Then $q_w = q_v^f$ and $\pi_v \in (\pi_w^e) \setminus (\pi_w^{e+1})$, so we see that

$$\log |\pi_v|_w = \log |\pi_w^e|_w = -e \log q_w = -ef \log q_v = r \log |\pi_v|_v$$

and consequently

$$\langle \cdot, \cdot \rangle_w = r \langle \cdot, \cdot \rangle_v.$$

The term "height" refers to the assignment $D \mapsto \lambda_{D,P_0}$; this can be viewed as a sort of height machine.

Proof of Theorem 4.12 (nonarchimedean). We will start by exhibiting a pairing for $D \in \mathrm{Div}^0(X)(F_v), E \in Z^0(X(F_v))$. Let $\mathcal{X}/\mathcal{O}_v$ a proper regular model for X. Let $\mathbf{D}, \mathbf{E}$ the Zariski closures. By Corollary 4.10 there is a $\mathbf{Q}$-divisor $\mathbf{D}'$ such that $\mathbf{D}' - \mathbf{D}$ is fibral and $\mathbf{D}' \cdot F_i = 0$ for all i (furthermore $\mathbf{D}'$ is unique up to $\mathbf{Q}\mathcal{X}_{\mathfrak{m}}$). Likewise we obtain an $\mathbf{E}'$. We set

$$\langle D, E \rangle_v \coloneqq -(\mathbf{D}' \cdot \mathbf{E}') \log q_v$$

Write $E = \sum n_i p_i$ for $p_i \in X(F_v)$ so that $\mathbf{E} = \sum n_i \mathbf{p}_i$. Suppose $\mathbf{D}'$ and $\mathbf{E}$ intersect on the special fiber at $x \in \mathcal{X}_{\mathfrak{m}}$, and let p be a point in the support of E so that $\mathbf{p}$ contains x. In a neighborhood of p, we can assume $\mathbf{D}'$ is principal and is cut out by $s \in K(X)$. Then

$$i_x(\mathbf{D}', \mathbf{p}) = \mathrm{length}_{\mathcal{O}_{\mathcal{X},x}} \mathcal{O}_{\mathcal{X},x}/((s) + \mathcal{I}_{\mathbf{p},x}) = \mathrm{length}_{\mathcal{O}_{\mathcal{X},x}} \mathcal{O}_{\mathbf{p},x}/s(\mathbf{p}) = \mathrm{length}_R R/s(\mathbf{p}) = v(s(p))$$

When $D = (f)$ is principal, then $\mathbf{D} = (f)$ (now the principal divisor on $\mathcal{X}$), and the above considerations show that

$$(\mathbf{D} \cdot \mathbf{E}) = \sum n_i v(f(p_i)) = -\log_{q_v} |f(E)|_v$$

Since $\mathbf{D} = \mathbf{D}'$ in this case, we have demonstrated that $\langle \cdot, \cdot \rangle_v$ has the desired behavior for principal divisors (it is fine that we used $\mathbf{E}$ rather than $\mathbf{E}'$ in our calculation, since $\mathbf{D}'$ cannot detect fibral divisors).

The remaining properties of $\langle \cdot, \cdot \rangle_v$ are easy to check and are left as an exercise.

To extend this to $E \in \mathrm{Div}^0(X)(F_v)$, consider an extension L_w/F_v over which $E \in Z^0(X(L_w))$ and apply the previous construction (renormalizing as necessary).

To show uniqueness, observe that the difference of any two such pairings descends to a pairing $J(F_v) \times J(F_v) \to \mathbf{R}$, continuous in each variable. Such a pairing must vanish since $J(F_v)$ is compact. □

Proof of Theorem 4.12 (archimedean). To be added. □

Theorem 4.13. *Let F a number field and X/F a smooth curve. Let $\langle \cdot, \cdot \rangle_F$ denote the Néron–Tate canonical height pairing on $J(F)$ associated to 2Θ. Then*

$$\langle \cdot, \cdot \rangle_F = \sum_v \langle \cdot, \cdot \rangle_v$$

for pairs of relatively prime divisors.

Proving this decomposition theorem would take us too far afield (plus I do not know where a proof exists); we accept it on faith.

5 Non-archimedean local heights for $X_0(N)$

Recall that we have so far reduced the main theorem to a statement relating $L'_{\mathscr{A}}(-, 1)$ to the series $g_{\mathscr{A}}$ of Theorem 3.9, given explicitly as

$$g_{\mathscr{A}} = \sum_{m \geq 1} \langle c, T_m c^\sigma \rangle q^m$$

To study this q-series, we need to unravel the global height pairings $\langle c, T_m c^\sigma \rangle$ in terms of local height pairings. This occurs in [5, §III] and the deformation-theoretic arguments that appear are carried out in detail in [1].

5.1 Reducing to intersection products

We want to compute the quantities $\langle c, T_m c^\sigma \rangle_v$ for each place v of H.

Fix a place v of H, let H_v the completion of H at v, with residue field k_v. Via Drinfeld level structures, we always have a proper model $\mathcal{X}/\mathcal{O}_v$ of the modular curve $X = X_0(N)$. For ease, we will assume we are in a situation where this model is regular[2] (for example when $v \nmid N$, or when $v \mid N$ and $j = 0, 1728$ are not supersingular).

Lemma 5.1. *Let $d = (x) - (0) \in \mathrm{Div}^0(X)(H)$. Then $\langle c, T_m c^\sigma \rangle = \langle c, T_m d^\sigma \rangle$.*

Proof. We have $c - d = (0) - (\infty)$, which is a torsion point in $J(H)$ (for example, by the Manin–Drinfeld theorem). □

As such, we can just compute $\langle c, T_m d^\sigma \rangle$ (this is useful because it gets rid of the common ∞ in the supports of our divisors). Let $\mathbf{c}$ and $\mathbf{d}$ denote the extensions to divisors on $\mathcal{X}$ via Zariski closure. Let p be the prime of $\mathbf{Z}$ under v. Write $N = p^t r$ where $(r, N) = 1$.

Proposition 5.2. *Either $\mathbf{c}$ or $\mathbf{d}$ intersects every fibral divisor trivially.*

We will need the following lemma.

Lemma 5.3. *Suppose p splits in $\mathcal{O}_K$. Let $\mathcal{E}/\mathcal{O}_{H_v}$ be an elliptic curve with complex multiplication by $\mathcal{O}_K$. Let $\mathfrak{p}$ be the prime of $\mathcal{O}_K$ under v and choose $\beta \in \mathcal{O}_K - \mathfrak{p}$. Then $[\beta] : \mathcal{E} \to \mathcal{E}$ is étale.*

Proof. The module $H^0(\mathcal{E}, \Omega^1_{\mathcal{E}/\mathcal{O}_{H_v}})$ is free of rank 1 over $\mathcal{O}_{H_v}$. We can fix an embedding $\mathcal{O}_{K_\mathfrak{p}} \hookrightarrow \mathcal{O}_{H_v}$ so that $\alpha \in \mathcal{O}_K$ acts on 1-forms by multiplication by α. After reducing modulo v, the module $H^0(\mathcal{E}_{k_v}, \Omega^1_{\mathcal{E}_{k_v}/k_v})$ is free of rank 1 over k_v and has an action by $\mathcal{O}_K/\mathfrak{p}$. So β acts by a unit on global 1-forms modulo v.

On the other hand, if β is not étale, then $[\beta]$ is inseparable and so acts by zero on 1-forms. □

Proof of Proposition 5.2. If $p \nmid N$ the conclusion is clear, so assume $p \mid N$. Then $\mathcal{X}_\mathfrak{m}$ is the disjoint union with crossings of components $\mathcal{X}_{a,b} \simeq X_0(r)_{k_v}$ parametrizing cyclic (a, b)-isogenies. Suppose $\mathbf{x}$ is a Heegner section representing a cyclic isogeny $\phi : \mathcal{E} \to \mathcal{E}'$ of elliptic curves over $\mathcal{O}_{H_v}$. Let $\mathfrak{n}$ denote the ideal of $\mathcal{O}_K$ annihilating $\ker \phi$, so that $N = \mathfrak{n}\overline{\mathfrak{n}}$ and v divides exactly one of $\mathfrak{n}$ or $\overline{\mathfrak{n}}$.

If $v \mid \overline{\mathfrak{n}}$, then ϕ is étale by Lemma 5.3 and consequently $\mathbf{x}$ reduces to a non-crossings point on $\mathcal{X}_{0,n}$. In this case, since $\mathbf{0}$ also reduces to a point on $\mathcal{X}_{0,n}$, we see that $\mathbf{d}$ has zero intersection with each fibral component.

If $v \mid \mathfrak{n}$, then $\widehat{\phi}$ is étale, so $\mathbf{x}$ reduces to a point on $\mathcal{X}_{n,0}$ and we similarly see that $\mathbf{c}$ has zero intersection with each fibral component. □

For the remainder of this section, we assume that $r_{\mathscr{A}}(m) = 0$ and $(m, N) = 1$.

Corollary 5.4. *We have $\langle c, T_m d^\sigma \rangle_v = -(\mathbf{c} \cdot T_m \mathbf{d}^\sigma) \log q_v$.*

[2]The situation where the model is not regular is treated by taking a regular resolution, which does not affect the cusps or the Heegner points.

So our question is reduced to computing $(\mathbf{c} \cdot T_m \mathbf{d}^\sigma)$.

Lemma 5.5. *We have* $(\mathbf{c} \cdot T_m \mathbf{d}^\sigma) = (\mathbf{x} \cdot T_m \mathbf{x}^\sigma)$.

Proof. Expanding out definitions we have

$$(\mathbf{c} \cdot T_m \mathbf{d}^\sigma) = (\mathbf{x} \cdot T_m \mathbf{x}^\sigma) - (\boldsymbol{\infty} \cdot T_m \mathbf{x}^\sigma) - (\mathbf{x} \cdot \mathbf{0}) + (\boldsymbol{\infty} \cdot T_m \mathbf{0}).$$

The last three terms all vanish because the pairs of divisors in question are disjoint. □

5.2 Intersection products in terms of deformation theory

We observe that all intersection products can be computed after a base change to $W = W(\overline{k_v})$. For ease write $k = \overline{k}_v$. We now write $\mathcal{X}/W$ for the proper regular arithmetic surface over W corresponding to the $\mathcal{X}$ from the previous section. The following key result transforms our questions about intersection products to questions about deformation theory.

Theorem 5.6. *Let* $\mathbf{x}, \mathbf{y} \in \mathcal{X}(W)$ *any two non-cuspidal horizontal sections on* $\mathcal{X}$ *intersecting properly at a point* z *on the special fiber, and assume that* $\mathrm{Aut}(z) = \{\pm 1\}$. *Then* $(\mathbf{x} \cdot \mathbf{y})$ *may be computed as the minimal value* $n_0 \geq 0$ *such that* $\mathbf{y}$ *and* $\mathbf{x}$ *are not isomorphic over* W_{n_0}.

Proof. Because $\mathrm{Aut}(z) = \{\pm 1\}$, the completed local ring $\widehat{\mathcal{O}}_{\mathcal{X},z}$ is the universal deformation space for z to local artinian W-algebras (the completed local ring on the moduli stack is always the universal deformation space; the condition on the automorphism group ensures that the local ring on the coarse space is canonically isomorphic). By [6, Chapter 5], $\widehat{\mathcal{O}}_{\mathcal{X},z}$ is a complete 2-dimensional regular local W-algebra; since z lies in the relative smooth locus of $\mathcal{X}/W$ (by Proposition 4.2), we see that $\widehat{\mathcal{O}}_{\mathcal{X},z} \simeq W[\![T]\!]$. Since $\mathbf{x}, \mathbf{y}$ are deformations of z to W, they correspond to W-valued points on the universal deformation space, i.e. a W-valued choice of T. Let these choices be $T_{\mathbf{x}}, T_{\mathbf{y}} \in W[\![T]\!]$. We are now just computing the intersection $(T_{\mathbf{x}} \cdot T_{\mathbf{y}})$ on the formal disc $\operatorname{Spec} W[\![T]\!]$ over W. This intersection number is evidently the maximal m for which $T_{\mathbf{x}} \equiv T_{\mathbf{y}} \mod \pi^m$, which in turn is equivalent to the specified n_0. □

Theorem 5.7. *Let* $\mathbf{x} \in \mathcal{X}(W)$ *be a Heegner section. Then*

$$(\mathbf{x} \cdot T_m \mathbf{x}^\sigma) = \frac{1}{2} \sum_{n \geq 0} |\mathrm{Hom}_{W_n}(\mathbf{x}^\sigma, \mathbf{x})_{\deg m}|$$

where $\mathrm{Hom}_{W_n}(-,-)_{\deg m}$ *denotes degree* m *morphisms of elliptic curves over* W_n.

Proof when $p \nmid m$. The assumption $r_{\mathscr{A}}(m) = 0$ ensures that $\mathbf{x}$ and $T_m \mathbf{x}^\sigma$ are relatively prime on the generic fiber. Write $T_m \mathbf{x}^\sigma = \sum_C \mathbf{x}_C^\sigma$. We claim that there is a natural bijection

$$\mathrm{Hom}_{W_n}(\mathbf{x}^\sigma, \mathbf{x})_{\deg m} \simeq \bigsqcup_C \mathrm{Isom}_{W_n}(\mathbf{x}_C^\sigma, \mathbf{x}).$$

Given an isogeny $\mathbf{x}^\sigma \to \mathbf{x}$ over W_n of degree m, its kernel is a group scheme $\overline{C}$ of order m which is then automatically étale. It therefore uniquely lifts to some an étale subgroup scheme $C \subset \mathbf{x}^\sigma$ over W, and there is a corresponding isomorphism $\mathbf{x}_C^\sigma \to \mathbf{x}$ over W_n.

The result now follows from Theorem 5.6. □

The arguments will necessarily be more complicated when $p \mid m$. To address this case, we need the theory of canonical and quasi-canonical liftings. We will treat the theory of canonical liftings, and for the moment omit the theory of quasi-canonical liftings.

5.3 Deformation theory of ordinary elliptic curves

For context we recall some general results about the deformation theory of elliptic curves. The discussion will be simplified by the assumption that we are working over the algebraically closed field k, but many of the results work more generally.

Throughout we let R denote a complete noetherian local domain with algebraically closed residue field k as before (we will apply our results to the context where R is a finite integrally closed extensions of W). Let $F = \mathrm{Frac}(R)$.

Definition 5.8. Let $\mathrm{Def}(R)$ denote the category of triples

$$(E/k, \Gamma/R, \varepsilon : \Gamma_k \xrightarrow{\sim} E(p^\infty))$$

where E/k is an elliptic curve, Γ/R is a p-divisible group, and ε is an isomorphism of p-divisible groups over k.

Theorem 5.9 (Serre–Tate)**.** *The functor* $\mathbf{Ell}(R) \to \mathrm{Def}(R)$ *given by*

$$\mathcal{E}/R \mapsto (\mathcal{E}_k, \mathcal{E}(p^\infty), \varepsilon)$$

(where ε is the natural choice of isomorphism) is an equivalence of categories.

Theorem 5.10 (Tate)**.** *The p-adic Tate module functor*

$$\begin{array}{ccc} \left\{ \begin{array}{c} p\text{-divisible} \\ \text{groups over } R \end{array} \right\} & \to & \left\{ \begin{array}{c} \mathbf{Z}_p\text{-lattices with} \\ G_F\text{-action} \end{array} \right\} \\ \Gamma & \mapsto & T(\Gamma)(\overline{F}) \end{array}$$

is fully faithful.

We now turn to the specific case of ordinary elliptic curves. Recall that an elliptic curve E/k is ordinary iff its p-divisible group $\Gamma_0 \coloneqq E(p^\infty)$ has connected étale sequence

$$0 \longrightarrow \Gamma_0^0 \longrightarrow \Gamma_0 \longrightarrow \Gamma_0^{\text{ét}} \longrightarrow 0$$

where $\Gamma_0^0 \simeq \mu_{p^\infty}$ and $\Gamma_0^{\text{ét}} \simeq \mathbf{Q}_p/\mathbf{Z}_p$. Furthermore, this exact sequence is uniquely split.

Let Γ/R be any lift of Γ_0. Examining its connected étale sequence, we see that we have

$$0 \longrightarrow \Gamma^0 \longrightarrow \Gamma \longrightarrow \Gamma^{\text{ét}} \longrightarrow 0 \tag{5.1}$$

where necessarily $\Gamma^0 \simeq \mu_{p^\infty}$ and $\Gamma^{\text{ét}} \simeq \mathbf{Q}_p/\mathbf{Z}_p$.

Definition 5.11. Let E/k an ordinary elliptic curve. A deformation $\mathcal{E}/R$ of E is a *(Serre–Tate) canonical lift* of E if $\mathcal{E}(p^\infty)$ has split connected-étale sequence, i.e. is the trivial extension class in (5.1).

Given an elliptic curve $\mathcal{E}/R$, we will say that $\mathcal{E}$ is a *(Serre–Tate) canonical lift* if $\mathcal{E}_k$ is ordinary and $\mathcal{E}$ is a Serre–Tate canonical lift of $\mathcal{E}_k$.

The splittings are always unique as there are no morphisms $\Gamma^{\text{ét}} \to \Gamma^0$.

Lemma 5.12. *Let $E, E'/k$ be ordinary elliptic curves with Serre–Tate canonical lifts $\mathcal{E}, \mathcal{E}'/R$. Then the natural map*

$$\mathrm{Hom}_R(\mathcal{E}, \mathcal{E}') \to \mathrm{Hom}_k(E, E')$$

is a bijection.

Proof. By the Serre–Tate theorem, to give a map $\mathcal{E} \to \mathcal{E}'$ over R is the same as giving a map $E \to E'$ and a compatible map $\mathcal{E}(p^\infty) \to \mathcal{E}'(p^\infty)$ over R. Every endomorphism of $\mu_{p^\infty} \times \mathbf{Q}_p/\mathbf{Z}_p$ over k lifts uniquely over R, so we're done. □

Lemma 5.13. *Let $\mathcal{E}/R$ be an elliptic curve with complex multiplication by $\mathcal{O}_K$ and of ordinary reduction. Then $\mathcal{E}$ is a canonical lift.*

Proof. Let $\Gamma = \mathcal{E}(p^\infty)$ denote the associated p-divisible group over R. By Theorem 5.10, it will suffice to show that $T(\Gamma)(\overline{F}) \cong \mathbf{Z}_p \oplus \mathbf{Z}_p(1)$. We know already from the connected étale sequence that $T(\Gamma)(\overline{F})$ is an extension of $\mathbf{Z}_p(1)$ by $\mathbf{Z}_p$.

To show the extension is split, we make use of the CM action. We observe that $T(\Gamma)(\overline{F})$ is a free rank-1 module over $\mathbf{Z}_p \otimes \mathcal{O}_K \cong \mathbf{Z}_p \times \mathbf{Z}_p$ (since p splits in $\mathcal{O}_K$) compatibly with the G_F-action, since the G_F-action commutes with the CM action. So $T(\Gamma)(\overline{F})$ is a sum of two G_F-characters. Since we already know it is an extension of $\mathbf{Z}_p(1)$ by $\mathbf{Z}_p$, we conclude that this extension is split. □

Remark. We didn't really need to pass to Tate modules, we could also argue directly in terms of the p-divisible groups.

Now we can treat our intersection products for primes $p \mid m$ such that p splits in K. Note that the assumption $p \mid m$ implies $p \nmid N$, so all level structures are étale.

Proof of Theorem 5.7 when $p \mid m$ and p splits in K. We claim that both sides vanish. On the RHS (deformation theory side) it suffices to show that $\mathrm{Hom}_k(\mathbf{x}^\sigma, \mathbf{x})_{\deg m} = 0$. By Lemma 5.12 this is equivalent to $\mathrm{Hom}_W(\mathbf{x}^\sigma, \mathbf{x})_{\deg m} = 0$ which is true because $r_{\mathscr{A}}(m) = 0$.

On the LHS (intersection theory side) we need to show that $(\mathbf{x} \cdot T_m \mathbf{x}^\sigma) = 0$. Write $m = p^t r$ where $(r, p) = 1$. Since r is prime to p, we can write $T_r \mathbf{x}^\sigma$ as a sum of irreducible divisors $\mathbf{z}_i$ defined over W. It will suffice to show that for any $\mathbf{z} = \mathbf{z}_i$, we have $(\mathbf{x} \cdot T_{p^t} \mathbf{z}) = 0$.

We work over a finite integrally closed extension W'/W so that $T_{p^t}\mathbf{z}$ splits as a sum of irreducible divisors $\mathbf{y}$ representing diagrams of elliptic curves over W', none of which are equal to $\mathbf{x}$. Suppose that $\mathbf{x}$ and $\mathbf{y}$ intersect on the special fiber; then there is a p^t-isogeny $\mathbf{z}_k \to \mathbf{x}_k$ on special fibers which lifts uniquely to a p^t-isogeny $\mathbf{z} \to \mathbf{x}$ since $\mathbf{z}$ and $\mathbf{x}$ are each canonical lifts. This contradicts that $\mathbf{x}$ and $T_{p^t}\mathbf{z}$ didn't meet on the generic fiber. □

We will not give a proof of Theorem 5.7 for the supersingular cases, i.e. when $p \mid m$ and does not split in K.

5.4 Counting via quaternion algebras

In light of Theorem 5.7 we are reduced to computing the quantities

$$\left|\mathrm{Hom}_{W_n}(\mathbf{x}^\sigma, \mathbf{x})_{\deg m}\right|.$$

We have already seen that this vanishes when p is split in K, so we can focus on the case when p is not split.

In this situation, $\mathbf{x}$ has supersingular reduction and $R \coloneqq \mathrm{End}_k(\mathbf{x})$ is an order in a quaternion algebra $B/\mathbf{Q}$ which is nonsplit exactly at p and ∞. One has the following more refined statement:

Proposition 5.14. *With setup as above, $R_p \coloneqq R \otimes \mathbf{Z}_p$ is a maximal order in $B_p \coloneqq B \otimes \mathbf{Q}_p$. At all primes $\ell \neq p$, the order $R \otimes \mathbf{Z}_\ell$ is conjugate in $B \otimes \mathbf{Q}_\ell \simeq M_2(\mathbf{Q}_\ell)$ to the order*

$$\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbf{Z}_\ell) : c \equiv 0 \mod N \right\}.$$

Since $\mathbf{x}$ has CM by $\mathcal{O}_K$, we obtain an embedding $\mathcal{O}_K \hookrightarrow R$ and therefore also $K \hookrightarrow B$. We can write $B = \mathbf{Q}[\alpha, \beta]$ subject to $\alpha \in K$, $\alpha^2 = d \in \mathbf{Z}$, $\beta^2 = e \in \mathbf{Z}$, $\beta \notin K$, and $\alpha\beta = -\beta\alpha$. Now α acts as an involution on B by conjugation; its decomposition into ± 1 eigenspaces is

$$B = B_+ \oplus B_-$$

where $B = K$ and $B_- = K\beta$. Recall the reduced norm

$$\mathrm{N}(p + q\alpha + r\beta + s\alpha\beta) = p^2 - dq^2 - er^2 + des^2,$$

and one sees easily that according to the above decomposition $b = b_+ + b_-$ we have

$$\mathrm{N}(b) = \mathrm{N}(b_+) + \mathrm{N}(b_-).$$

Lemma 5.15. *We have a canonical identification*

$$\mathrm{End}_{W_n}(\mathbf{x}) \simeq \mathcal{O}_K + p^n R$$

compatible with the identification $\mathrm{End}_k(\mathbf{x}) \simeq R$.

Giving a full proof would take too much space, but let us do a quick sanity check. When $n = 0$ we get $\mathrm{End}_k(\mathbf{x}) = R$ and as $n \to \infty$ we recover $\mathrm{End}_W(\mathbf{x}) = \mathcal{O}_K$.

Corollary 5.16. *Suppose p is nonsplit in K. Then there is a canonical identification*

$$\mathrm{End}_{W_n}(\mathbf{x}) \simeq \{b \in R : \mathrm{N}(b_-) \equiv 0 \mod p^{\varepsilon(n)}\}, \quad \varepsilon(n) = \begin{cases} 2n+1, & p \text{ is inert} \\ n, & p \text{ is ramified} \end{cases}$$

Proof. Follows from Lemma 5.15. □

Proposition 5.17. *We have a canonical identification*

$$\mathrm{Hom}_{W_n}(\mathbf{x}^\sigma, \mathbf{x}) \simeq \mathrm{End}_{W_n}(\mathbf{x}) \cdot \mathfrak{a} \subseteq B$$

where $\mathfrak{a} \in \mathscr{A}$ is any representative. For any isogeny ϕ on the LHS corresponding to $b \in B$, we have $\deg \phi = \mathrm{N}(b)/N(\mathfrak{a})$.

We will omit the explanation of how to justify this over an arbitrary base; let us discuss the situation over H. Given a CM elliptic curve E/H with complex uniformization $\mathbf{C}/\mathfrak{b}$, the elliptic curve E^σ is given by $\mathbf{C}/\mathfrak{a}^{-1}\mathfrak{b}$ and an isogeny $E^\sigma \to E$ is determined by an inclusion of $\mathcal{O}$-modules $\mathfrak{a}^{-1}\mathfrak{b} \hookrightarrow \mathfrak{b}$, or equivalently an inclusion $\mathcal{O} \hookrightarrow \mathfrak{a}$, i.e. an element of $\mathfrak{a}$. So we get an identification $\mathrm{Hom}_H(E^\sigma, E) \cong \mathfrak{a}$, and the degree of an isogeny corresponding to $\gamma \in \mathfrak{a}$ is given by $N(\gamma)/N(\mathfrak{a})$.

Corollary 5.18. *We have*

$$\sum_{n\geq 0} \mathrm{Hom}_{W_n}(\mathbf{x}^\sigma, \mathbf{x})_{\deg m} = \sum_{\substack{b \in R\mathfrak{a}/\pm 1 \\ \mathrm{N}(b)/N(\mathfrak{a}) = m}} \begin{cases} \frac{1}{2}(1 + \mathrm{ord}_p(\mathrm{N}(b_-))), & p \text{ is inert} \\ \mathrm{ord}_p(\mathrm{N}(b_-)), & p \text{ is ramified} \end{cases}.$$

It is desirable to remove the reference to the quaternion algebra B and express the intersection number $(\mathbf{x} \cdot T_m \mathbf{x}^\sigma)$ entirely in terms of K. To do this, we provide an explicit realization of the algebra B. First assume p is inert, and choose a prime q such that $q \equiv -p \bmod D$. Then we can take B to be the quaternion algebra with $\alpha = \sqrt{D}$, $\beta = \sqrt{-pq}$.

By computing explicitly in this algebra, one finds the following formula:

Proposition 5.19. *If p is inert in K, then*

$$\langle c, T_m d^\sigma \rangle_p := \sum_{v|p} \langle c, T_m d^\sigma \rangle_v = -(\log p) \sum_{n=1}^{\lfloor \frac{m|D|}{Np} \rfloor} 2^{\omega((n,D))} (2 + \mathrm{ord}_p(n)) r_{\mathscr{A}}(m|D| - pnN) r_{\mathscr{A}\mathfrak{q}\mathfrak{n}}(n)$$

where $\omega(-)$ denotes number of distinct prime factors.

For the details, see [5, §III.9]. We get a similar formula in the case that p is ramified.

References

[1] B. Conrad. Gross–Zagier revisited, 2010.

[2] P. Deligne and M. Rapoport. Les schèmas de modules de courbes elliptiques, 1973.

[3] B. Gross. Heegner points on $X_0(N)$, 1984.

[4] B. Gross. Local heights on curves, 1986.

[5] B. Gross and D. Zagier. Heegner points and derivatives of L-series, 1986.

[6] N. Katz and B. Mazur. Arithmetic moduli of elliptic curves, 1985.

[7] S. Lang. Introduction to Arakelov Theory, 1988.

[8] J. Silverman. Advanced Topics in the Arithmetic of Elliptic Curves, 1994.