p-adic modular forms

TCC (Spring 2021), Lecture 6

Pak-Hin Lee

25th February 2021

Administrative issues

Slides:

- Lectures 1-5: available on webpage
- Lecture 5, P.23: My generalization of lemme 3 isn't quite right. See Problem Sheet 2.

Problem sheets:

- 3 sets for assessment
  1. 22nd February (Monday of Week 6): posted
  2. 8th March (Monday of Week 8): posted
  3. 22nd March (Monday of Week 10): tentative
- available at least two weeks before deadlines

Administrative issues

Office hours:

- Day: Tuesdays during Weeks 7, 8, 9 (tentative)
- Time: TBA
- Format: Q&A, possibly supplementary discussion or lectures
- Email: Please respond!

References for geometric modular forms:

- Calegari's AWS notes, 2013
- Loeffler's TCC notes, 2014
- Katz, §1 and Appendix 1
- I might go over some skipped details during office hours.

Plans

Today:

- Reinterpretation of modular forms over **C**
- Algebro-geometric interpretation of modular forms
- Tate curve and q-expansions

Next week:

- q-expansion principle
- Hasse invariants
- p-adic modular forms, finally!

Overview: Geometric modular forms

Goal

Interpret modular forms using algebraic geometry.

- Complex analysis: Modular forms are initially defined as holomorphic functions on $\mathbf{H}$ satisfying a transformation property.
- Lattices: Interpret as functions on lattices $\Lambda \subset \mathbf{C}$.
- Weierstrass parametrization: Interpret as functions on elliptic curves over $\mathbf{C}$ (with additional data).
- Algebraic geometry: Generalize this for elliptic curves over any ring (or scheme).

Weierstrass parametrization

Weierstrass parametrization: For a lattice $\Lambda \subset \mathbf{C}$, the complex torus $\mathbf{C}/\Lambda$ has the structure of an elliptic curve with equation

$$y^2 = 4x^3 - 60G_4(\Lambda)x - 140G_6(\Lambda)$$

where

$$G_{2k}(\Lambda) := \sum_{\lambda \in \Lambda - \{0\}} \frac{1}{\lambda^{2k}}.$$

The isomorphism is given by

$$x = \wp(z;\Lambda) := \frac{1}{z^2} + \sum_{\lambda \in \Lambda - \{0\}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right),$$

$$y = \wp'(z;\Lambda) = -\sum_{\lambda \in \Lambda} \frac{2}{(z-\lambda)^3}.$$

Weierstrass parametrization

- Homothety: Two lattice Λ and Λ' are homothetic, denoted $\Lambda \sim \Lambda'$ if $\Lambda = \mu\Lambda'$ for some $\mu \in \mathbf{C}^\times$.
- Two homothetic lattices give rise to isomorphic elliptic curves, and vice versa:

$$\Lambda \sim \Lambda' \iff \mathbf{C}/\Lambda \cong \mathbf{C}/\Lambda'.$$

- Weierstrass parametrization: There is a bijection

$$\{\text{Lattices in } \mathbf{C}\}/\sim \longleftrightarrow \{\text{Elliptic curves over } \mathbf{C}\}/\cong$$
$$\Lambda \longmapsto \mathbf{C}/\Lambda.$$

Modular forms as functions on lattices

- Every lattice is homothetic to one of the form

$$\mathbf{Z}\tau + \mathbf{Z}, \quad \tau \in \mathbf{H}.$$

- We have

$$\mathbf{Z}\tau + \mathbf{Z} \sim \mathbf{Z}\tau' + \mathbf{Z} \iff \tau' = \frac{a\tau + b}{c\tau + d}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}).$$

- Modular forms of weight k can be interpreted as functions on lattices satisfying

$$F(\mu\Lambda) = \mu^{-k} F(\Lambda).$$

This correspondence is given by $f(\tau) = F(\mathbf{Z}\tau + \mathbf{Z})$.

Example

$G_{2k}(\mu\Lambda) = \mu^{-2k} G_{2k}(\Lambda)$.

Lattices and elliptic curves

Can we upgrade the following diagram?

$$\begin{array}{ccc}
\{\text{Lattices in } \mathbf{C}\} & \longleftrightarrow & ? \\
\downarrow & & \downarrow \\
\{\text{Lattices in } \mathbf{C}\}/\sim & \longleftrightarrow & \{\text{Elliptic curves over } \mathbf{C}\}/\cong
\end{array}$$

$$\Lambda \longmapsto \mathbf{C}/\Lambda : y^2 = 4x^3 - 60G_4(\Lambda)x - 140G_6(\Lambda)$$

Lemma

The space $H^0(E, \Omega^1)$ of holomorphic 1-forms on an elliptic curve E over $\mathbf{C}$ is one-dimensional.

For $E = \mathbf{C}/\Lambda$, $H^0(E, \Omega^1) = \mathbf{C} \cdot dz$.

Lattices and elliptic curves

- Thus there is a bijection

$$\{\text{Lattices in } \mathbf{C}\} \longleftrightarrow \left\{(E,\omega) : \omega \in H^0(E,\Omega^1) - \{0\}\right\} / \cong$$
$$\Lambda \longmapsto (\mathbf{C}/\Lambda, dz).$$

- If Λ corresponds to (E,ω), then $\mu\Lambda$ corresponds to $(E,\mu\omega)$.
- Modular forms of weight k can be interpreted as functions on (E,ω) satisfying

$$f(E,\mu\omega) = \mu^{-k} f(E,\omega)$$

(with some condition at ∞).

Notation

We want to generalize this for any ring R (in fact, any scheme S). Notation:

- R_0: base ring
- S: scheme over R_0
- E/S: elliptic curve over S, i.e. a morphism $p : E \to S$ such that p is smooth and proper and all fibers are connected curves of genus 1, together with a section $s : S \to E$
- Then $\underline{\omega}_{E/S} := p_*(\Omega^1_{E/S})$ is an invertible sheaf on S.

Elliptic curves over schemes

If this scares you...

Remark

Locally on $S = \operatorname{Spec} R$ (where R is an R_0-algebra):

- E is given by a Weierstrass equation over R;
- $\underline{\omega}_{E/R}$ is a free R-module of rank 1.

Thus everything can be made explicit!

For details, see [Loeffler, Proposition 3.3.2].

Geometric modular forms

Definition

A **modular form** of level 1 and weight k with coefficients in R_0 is a rule f which assigns

$$E/S \mapsto f(E/S) \in H^0(S, \underline{\omega}_{E/S}^{\otimes k})$$

with the following properties:

1. $f(E/S)$ depends only on the S-isomorphism class of E/S;
2. the formation of $f(E/S)$ commutes with base change, i.e. for any fiber diagram

$$\begin{array}{ccc} E_{S'} & \longrightarrow & S' \\ \downarrow & \square & \downarrow g \\ E & \longrightarrow & S \end{array}$$

   we have $f(E_{S'}/S') = g^*(f(E/S))$.

Geometric modular forms

- The R_0-module of such forms is denoted $M(R_0; 1, k)$.
- The test objects S vary over R_0-schemes, but it is enough to consider the affine ones.
- **Idea:** Base change and gluing.

Geometric modular forms

Equivalent definition: $f \in M(R_0; 1, k)$ is a rule which assigns to every pair $(E/R, \omega)$ where

- R is an R_0-algebra;
- E/R is an elliptic curve over R;
- ω is a basis of $\underline{\omega}_{E/R}$

an element

$$f(E/R, \omega) \in R,$$

with the following properties:

1. $f(E/R)$ depends only on the R-isomorphism class of $(E/R, \omega)$;
2. $f(E, \lambda\omega) = \lambda^{-k} f(E, \omega)$ for any $\lambda \in R^\times$;
3. the formation of $f(E/R)$ commutes with base change, i.e. for any ring map $\phi : R \to R'$, we have $f(E/R', \omega_{R'}) = \phi(f(E/R, \omega))$.

Geometric modular forms

Given the second definition, we recover the first definition by defining the section

$$f(E/R,\omega)\omega^{\otimes k} \in \underline{\omega}_{E/R}^{\otimes k}$$

which is independent of the choice of basis ω by (2).

Holomorphicity

Holomorphicity:

- The base change condition captures “continuity” and even “holomorphicity”.
- **Idea:** Two test objects that are “close” can be put in a family.

Holomorphicity at ∞:

- What about “holomorphicity at ∞”? Note that we don’t have analysis!
- We will see that the base change condition guarantees “meromorphicity at ∞”, so $M(R_0; 1, k)$ can be thought of as the space of meromorphic modular forms.
- To understand behavior at ∞ (more precisely, q-expansions), we introduce a special test object: the Tate curve $\mathrm{Tate}(q)$.

Tate curve

Definition

The **Tate curve** is the elliptic curve $\mathrm{Tate}(q)$ over $\mathbf{Z}((q)) := \mathbf{Z}[[q]][\frac{1}{q}]$ defined by

$$y^2 + xy = x^3 + a_4(q)x + a_6(q)$$

where

$$a_4(q) := -5S_3(q),$$
$$a_6(q) := -\frac{1}{12}\left(5S_3(q) + 7S_5(q)\right),$$
$$S_k(q) := \sum_{n=1}^{\infty} \sigma_k(n) q^n \in \mathbf{Z}[[q]].$$

Tate curve

Remark

Tate(q) has discriminant given by the normalized weight 12 cusp form

$$\Delta = q \prod_{n=1}^{\infty} (1 - q^n)^{24}$$

and j-invariant given by the j-function

$$j = q^{-1} + 744 + 196884q + \cdots .$$

In particular, this explains why Tate(q) is an elliptic curve over $\mathbf{Z}((q))$ – although its coefficients lie in $\mathbf{Z}[[q]]$, one has to invert Δ:

$$\mathbf{Z}[[q]][\Delta^{-1}] = \mathbf{Z}((q)).$$

Tate curve

Where does this come from?

- By a change of variables, the equation can be rewritten as

$$Y^2 = 4X^3 - \frac{1}{12}E_4(q)X + \frac{1}{216}E_6(q)$$

  where $E_k(q) = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n$ is (the q-expansion of) the Eisenstein series. This is only defined over $\mathbf{Z}[\frac{1}{6}]$, but the previous equation always has $\mathbf{Z}$-coefficients.
- Let us interpret how this equation arises (**bonus:** canonical differential ω_{can} on $\mathrm{Tate}(q)/\mathbf{Z}((q))$).
- For simplicity, we work over $\mathbf{Z}[\frac{1}{6}]$.

Tate curve

- Given a lattice Λ, recall the Weierstrass parametrization

$$\begin{aligned} \mathbf{C}/\Lambda &\to \mathbf{P}^2 \\ z &\mapsto [\wp(z;\Lambda), \wp'(z;\Lambda), 1] \end{aligned}$$

where $\wp(z;\Lambda)$ is the Weierstrass $\wp$-function

$$\wp(z;\Lambda) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda-\{0\}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right).$$

- Consider the lattice $\Lambda_\tau := \mathbf{Z} + \mathbf{Z}\tau$, $\tau \in \mathbf{H}$. The exponential map $e^{2\pi i -} : \mathbf{C}/\mathbf{Z} \xrightarrow{\sim} \mathbf{C}^\times$ induces an isomorphism

$$\begin{aligned} \mathbf{C}/\Lambda_\tau &\xrightarrow{\sim} \mathbf{C}^\times/q^{\mathbf{Z}} \\ z &\mapsto e^{2\pi i z} \end{aligned}$$

where $q = e^{2\pi i\tau}$.

Tate curve

- In terms of the parameter $u = e^{2\pi i z}$ on $\mathbf{C}^\times/q^{\mathbf{Z}}$, we rewrite

$$\wp(z;\Lambda_\tau) = F(u;q),$$
$$\wp'(z;\Lambda_\tau) = G(u;q).$$

  Then F and G define an isomorphism $\mathbf{C}^\times/q^{\mathbf{Z}} \cong \mathrm{Tate}(q)$.
- Transporting the canonical differential dz on $\mathbf{C}/\Lambda_\tau$ to $\mathbf{C}^\times/q^{\mathbf{Z}}$ gives:

Definition

The **canonical differential** ω_{can} on the Tate curve

$$\mathrm{Tate}(q) : y^2 + xy = x^3 + a_4(q)x + a_6(q)$$

is

$$\omega_{\mathrm{can}} := \frac{dx}{2y+x}.$$

Tate curve

- For details, see [Katz, Appendix 1] and carry out the calculations.
- Over $\mathbf{C}$, every elliptic curve is parametrized by (a specialization of) the Tate curve; this is essentially by construction (and Weierstrass parametrization).
- In general, this is not true.

Example (Tate)

Over a p-adic field K, the formal power series involved in Tate(q) turn out to be **convergent** for $0 < |q| < 1$. Then Tate(q) is an elliptic curve with $|j| > 1$, and we can identify

$$\overline{K}^{\times}/q^{\mathbf{Z}} \cong \mathrm{Tate}(q)(\overline{K})$$

via F and G. Furthermore, this isomorphism is Galois-equivariant.

Digression: Tate uniformization

In particular, any elliptic curve E/K with $|j(E)| \leq 1$ **cannot** be parametrized by the Tate curve.

Theorem (Tate)

Let K be a finite extension of $\mathbf{Q}_p$.

1. *Given an elliptic curve E/K with $|j(E)| > 1$, there exists a unique $q \in K^\times$ with $|q| < 1$ such that $E \cong \mathrm{Tate}(q)$ over $\overline{K}$.*
2. *This isomorphism descends to K if and only if E has split multiplicative reduction.*

Remark

- The isomorphism in (1) always descends to a quadratic extension of K.
- This result is what led Tate to introduce the Tate curve and subsequently develop the theory of rigid analytic geometry.

q-expansions

Idea

- Understand Tate(q) as a family of elliptic curves over the punctured disk.
- The behavior at $q = 0$ tells us how the curve degenerates at ∞.

Let $f \in M(R_0; 1, k)$ be a modular form over R_0.

- Evaluating f at the Tate curve gives its **q-expansion**

$$f(\mathrm{Tate}(q), \omega_{\mathrm{can}}) \in \mathbf{Z}((q)) \otimes_{\mathbf{Z}} R_0.$$

- *A priori* this has a finite tail, so f can be thought of as being automatically **meromorphic at** ∞.

q-expansions

- f is said to be **holomorphic at** ∞ if

$$f(\mathrm{Tate}(q), \omega_{\mathrm{can}}) \in \mathbf{Z}[[q]] \otimes_{\mathbf{Z}} R_0,$$

and a **cusp form** if

$$f(\mathrm{Tate}(q), \omega_{\mathrm{can}}) \in q\mathbf{Z}[[q]] \otimes_{\mathbf{Z}} R_0.$$

- Warning: $\mathbf{Z}[[q]] \otimes_{\mathbf{Z}} R_0 \subsetneq R_0[[q]]$.
- The set of holomorphic modular forms is denoted

$$S(R_0; 1, k) \subset M(R_0; 1, k).$$

Remark (Notation)

Katz's notation differs from the classical usage, where S and M denote the cusp forms and holomorphic modular forms respectively.

Higher levels

- If there is a universal element in the moduli space of test objects E/S, then we can simply pull back.
- Unfortunately, there is no such element in the level 1 case.
- Now we generalize for higher levels.
- Fix a positive integer N, and S will denote a scheme over $\mathbf{Z}[\frac{1}{N}]$.

Remark

Katz–Mazur works with Drinfeld level structures, which allows working over $\mathbf{Z}$.

Higher levels

Definition

A level N structure for E/S is an isomorphism of group schemes

$$\alpha_N : E[N] \xrightarrow{\sim} (\mathbf{Z}/N\mathbf{Z})^2_S.$$

Remark

- For this to exist, N has to be invertible on S.
- If this exists (and S is connected), then the set of level N structures is a torsor for $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$.

Modular forms of higher levels

- The (meromorphic) modular forms of level N and weight k, denoted $M(R_0; N, k)$, are rules f

$$(E/S, \alpha_N) \mapsto f(E/S, \alpha_N) \in H^0(S, \underline{\omega}_{E/S}^{\otimes k})$$

or equivalently

$$(E/R, \omega, \alpha_N) \mapsto f(E/R, \omega, \alpha_N) \in R$$

satisfying the evident properties.
- To talk about q-expansions and holomorphicity at ∞, we need to study the Tate curve.

Tate curve at higher level

Definition

The **Tate curve at level** N is $\mathrm{Tate}(q^N)$ over $\mathbf{Z}((q))$ defined by

$$y^2 + xy = x^3 + a_4(q^N)x + a_6(q^N)$$

where a_4 and a_6 are as before. The canonical differential is

$$\omega_{\mathrm{can}} = \frac{dx}{2y + x}.$$

- Fix a primitive N-th root of unity ζ_N.
- The N-torsion of $\mathbf{C}^\times / q^{N\mathbf{Z}}$ is $\{\zeta_N^i q^j : 0 \leq i, j \leq N - 1\}$.
- On $\mathrm{Tate}(q^N)$, these are defined over $\mathbf{Z}[[q]] \otimes_{\mathbf{Z}} \mathbf{Z}[\zeta_N, \frac{1}{N}]$.

q-expansions at higher level

Suppose R_0 contains $\frac{1}{N}$ and ζ_N.

- For each level N structure α on $\mathrm{Tate}(q^N)$, the q**-expansion** of $f \in M(R_0; N, k)$ at α is

$$f(\mathrm{Tate}(q^N), \omega_{\mathrm{can}}, \alpha) \in \mathbf{Z}((q)) \otimes_{\mathbf{Z}} R_0.$$

- Again, meromorphicity at ∞ is automatic.
- f is **holomorphic at** ∞ (resp. **a cusp form**) if for all level N structures α, its q-expansion at α belongs to $\mathbf{Z}[[q]] \otimes_{\mathbf{Z}} R_0$ (resp. $q\mathbf{Z}[[q]] \otimes_{\mathbf{Z}} R_0$).
- The space of holomorphic forms is denoted $S(R_0; N, k) \subset M(R_0; N, k)$.

Modular curves

A brief summary:

- For $N \geq 3$, there exists a (fine) moduli scheme $Y(N)$ parametrizing elliptic curves with level N structure:

$$S \text{ scheme over } \mathbf{Z}[\frac{1}{N}] \rightsquigarrow \{(E/S, \alpha_N)\}/\sim .$$

- $Y(N)$ is a smooth affine curve over $\mathbf{Z}[\frac{1}{N}]$.
- Its "compactification" $X(N)$ is a smooth proper curve over $\mathbf{Z}[\frac{1}{N}]$.
- These come with universal elliptic curves

$$\begin{array}{ccc} \mathcal{E} & \longrightarrow & \overline{\mathcal{E}} \\ {\scriptstyle\pi}\downarrow & & \downarrow \\ Y(N) & \longrightarrow & X(N). \end{array}$$

Modular curves

- The invertible sheaf $\underline{\omega} := \pi_*(\Omega_{\mathcal{E}/Y(N)})$ on $Y(N)$ extends uniquely to $X(N)$ ("weight 1 modular forms").

Theorem (Consequence of Kodaira–Spencer isomorphism)

For any $\mathbf{Z}[\frac{1}{N}]$-algebra R_0, we have

$$S(R_0; N, k) = H^0(X(N), \underline{\omega}^{\otimes k} \otimes_{\mathbf{Z}[\frac{1}{N}]} R_0).$$

Definition

For any $\mathbf{Z}[\frac{1}{N}]$-module R_0, we define

$$S(R_0; N, k) := H^0(X(N), \underline{\omega}^{\otimes k} \otimes_{\mathbf{Z}[\frac{1}{N}]} R_0).$$

Here $\underline{\omega}^{\otimes k} \otimes_{\mathbf{Z}[\frac{1}{N}]} R_0$ is a quasi-coherent sheaf on $X(N)$.

Modular curves

- For level $N = 1$ or 2, $Y(N)$ only exists as a coarse moduli scheme.
- **Solution:** Use stacks or the following trick.
- Lift to a covering and descend back:

$$S(R_0; 1, k) := S(R_0; 3, k)^{\mathrm{GL}_2(\mathbf{Z}/3\mathbf{Z})}$$

and similarly for $S(R_0; 2, k)$.

Remark

Each statement in the summary above has an analogue in the complex-analytic setting.