Modular Curves

September 4, 2013

The first examples of Shimura varieties we encounter are the modular curves. In this lecture we review the basics of modular curves, beginning with the complex theory and progressing towards modular curves over number fields.

1 Modular curves as complex manifolds

1.1 Lattices and the upper half plane

$\mathcal{H}$ is the upper half plane, a complex manifold. It will be helpful to interpret $\mathcal{H}$ in multiple ways.

A *lattice* $\Lambda \subset \mathbb{C}$ is a free abelian group of rank 2, for which the map $\Lambda \otimes_{\mathbb{Z}} \mathbb{R} \to \mathbb{C}$ is an isomorphism. In other words, Λ is a subgroup of $\mathbb{C}$ of the form $\mathbb{Z}\alpha \oplus \mathbb{Z}\beta$, where $\{\alpha, \beta\}$ is basis for $\mathbb{C}/\mathbb{R}$. Two lattices Λ and Λ' are *homothetic* if $\Lambda' = \theta\Lambda$ for some $\theta \in \mathbb{C}^*$. This is an equivalence relation, and the equivalence classes are *homothety classes*.

Let's consider $\mathbb{C}$ as an *oriented* real vector space, meaning we have a privileged basis of $\bigwedge^2 \mathbb{C}$ modulo scaling by a positive real number. Then an *oriented basis* of a lattice Λ is a basis $\{a + bi, c + di\}$ with $ad - bc > 0$.

Any lattice Λ is homothetic to one of the form $\mathbb{Z} \oplus \mathbb{Z}\tau$, where $\tau \in \mathcal{H}$. The following is very easy:

Proposition 1.1.1. *The map $\tau \mapsto (\mathbb{Z} \oplus \mathbb{Z}\tau, \{1, \tau\})$ is a bijection between $\mathcal{H}$ and the set of homothety classes of pairs $(\Lambda, \{\alpha, \beta\})$, where $\Lambda \subset \mathbb{C}$ is a lattice and $\{\alpha, \beta\}$ is an oriented basis for Λ.*

If a lattice Λ has two oriented bases $\{\alpha, \beta\}$ and $\{\gamma, \delta\}$, then the change of basis matrix between them lies in $\mathrm{SL}_2(\mathbb{Z})$. The action of $\mathrm{SL}_2(\mathbb{Z})$ on the

set of oriented bases of a lattice corresponds to the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d}.$$

In light of this, $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ classifies the set of homothety classes of lattices in $\mathbb{C}$.

One has to be a little careful with the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ because two of the $\mathrm{SL}_2(\mathbb{Z})$-orbits in $\mathcal{H}$ have nontrivial stabilizer in $\mathrm{PSL}_2(\mathbb{Z})$. These are i and $e^{2\pi i/3}$, whose stabilizers in $\mathrm{PSL}_2(\mathbb{Z})$ have orders 2 and 3, respectively. With some care, it is possible to give $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ the structure of a complex manifold, rather than an orbifold.

Proposition 1.1.2. *The j-function*

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + \ldots, \; q = e^{2\pi i\tau}$$

gives an isomorphism of complex manifolds $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H} \to \mathbb{C}$.

1.2 A crash course in elliptic curves

Informally, there are (at least) three ways of looking at an elliptic curve:

- An elliptic curve is a smooth projective curve of genus 1 over a field K, together with a rational base point $O \in E(K)$.
- An elliptic curve is a smooth curve in projective space cut out by a Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$.
- An elliptic curve is a complex torus of dimension 1, equal to the quotient $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$.

Of these, the first is the most powerful definition. The first and second are equivalent, and the first and third are equivalent when the base field is $\mathbb{C}$. It might be helpful to (very quickly) review these equivalences.

Let E/K be a smooth projective curve of genus 1, and let $O \in E(K)$ be a rational point. Recall that a divisor D on a curve is a formal sum of points with $\mathbb{Z}$-coefficients, and if $D = \sum_P a_P[P]$ is a divisor, then $H(D)$ is the vector space of rational functions f on the curve which satisfy $\mathrm{ord}_P(f) \geq -a_P$ for

all P. We can use the Riemann-Roch theorem to compute the dimension of $H(n[O])$ for all $n \geq 0$: we have

$$\dim H(n[O]) = \begin{cases} 1, & n = 0, 1 \\ n, & n \geq 2. \end{cases}$$

This means that $H(0) = H([O]) = K$, $H(2[O]) = \langle 1, x \rangle$ for some rational function x with a double pole at O, and $H(3[O]) = \langle 1, x, y \rangle$ for some other rational function y with a triple pole at O. We have $\dim H(6[O]) = 6$, and $H(6[O])$ contains $1, x, x^2, x^3, y, xy, y^2$, which must therefore be linearly dependent. The equation of linear dependence among these functions determines a Weierstrass equation for E.

Conversely, if E/K is a smooth projective curve cut out by a Weierstrass equation, then E has genus $3(3-1)/2 = 1$, and the point at infinity is rational. Thus the first and second definitions are equivalent.

Now assume the base field is $\mathbb{C}$. If $E/\mathbb{C}$ is an elliptic curve, then $E(\mathbb{C})$ is a Riemann surface of genus 1, and therefore the space $H^0(E(\mathbb{C}), \Omega^1)$ is 1-dimensional. Let ω be a basis vector. On the other hand, $\Lambda := H_1(E(\mathbb{Z}), \mathbb{Z})$ is a free $\mathbb{Z}$-module of rank 2. We have a map $\Lambda \to \mathbb{C}$ given by $\gamma \mapsto \int_\gamma \omega$, which induces an isomorphism $\Lambda \otimes_{\mathbb{Z}} \mathbb{R} \cong \mathbb{C}$. Thus Λ is a lattice in $\mathbb{C}$, and we have an isomorphism of complex manifolds

$$\begin{aligned} E(C) &\to \mathbb{C}/\Lambda \\ P &\mapsto \int_O^P \omega. \end{aligned}$$

Conversely, if we are given a lattice $\Lambda \in \mathbb{C}$, we have the Weierstrass function $\wp_{\Lambda(z)}$, a Λ-periodic meromorphic function with a double pole at every point in Λ. Then $\wp_\Lambda$ satisfies a differential equation of the form

$$[\wp_\Lambda(z)']^2 = 4\wp_\Lambda(z)^3 - g_2\wp_\Lambda(z) - g_3$$

for constants g_2 and g_3 depending on Λ. The discriminant of the cubic polynomial on the right is nonzero, so that the above may be interpreted as a Weierstrass equation defining an elliptic curve $E/\mathbb{C}$. Then $z \mapsto (\wp_\Lambda(z), \wp'_\Lambda(z))$ is an isomorphism of complex manifolds between C/Λ and $E(\mathbb{C})$.

We can now give an interpretation of $\mathcal{H}$ in terms of elliptic curves. Note that if $E/\mathbb{C}$ is an elliptic curve, then $H_1(E(\mathbb{C}), \mathbb{R})$ is an oriented vector space (under the intersection pairing), so that it makes sense to talk about a basis for $H_1(E(\mathbb{C}), \mathbb{Z})$ being oriented.

Proposition 1.2.1. *$\mathcal{H}$ classifies isomorphism classes of pairs $(E, \{\alpha, \beta\})$, where $E/\mathbb{C}$ is an elliptic curve and $\{\alpha, \beta\}$ is an oriented basis for $H_1(E(\mathbb{C}), \mathbb{Z})$. The quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ classifies isomorphism classes of elliptic curves over $\mathbb{C}$.*

1.3 The j-line and the λ-line

Isomorphism classes of elliptic curves over $\mathbb{C}$ are in bijection with $\mathbb{C}$ itself. Is there an algebraic family of elliptic curves parametrized by the affine j-line $\mathbb{A}^1_j$, such that the fiber over $j = j_0$ is the elliptic curve with j-invariant j_0? This would mean a Weierstrass equation

$$y^2 = 4x^3 - g_2 x - g_3$$

with $g_2, g_3 \in \mathbb{C}[j]$, which defines an elliptic curve for all $j \in \mathbb{C}$, and whose j-invariant is j. This means that the discriminant Δ must have no roots in $\mathbb{C}$, *i.e.* it is a scalar. The j-invariant is $1728 g_2^3/\Delta$, which is a cube in $\mathbb{C}[j]$, so that it cannot equal j. Furthermore, $j - 1728 = 1728 \times 27 g_3^2/\Delta$, which means that $j - 1728$ must be a square in $\mathbb{C}[j]$, another contradiction.

The above phenomena are quite related to the fact that the elliptic curves of j-invariants $j(e^{2\pi/3}) = 0$ and $j(i) = 1728$ have extra automorphisms on top of the usual ± 1, by a factor of 3 and 2, respectively.

We can resolve this problem by passing to a 6-fold cover of the j-line. Consider the Weierstrass equation

$$y^2 = x(x-1)(x-\lambda),$$

which defines an elliptic curve E_λ for $\lambda \in \mathbb{P}^1 \backslash \{0, 1, \infty\}$ (the λ-line), together with a basis $\{(0,0), (1,0)\}$ for the 2-torsion of E_λ. It turns out that any elliptic curve over $\mathbb{C}$ equipped with a basis for its 2-torsion corresponds to a unique value of λ.

The j-invariant of E_λ is

$$j = 256 \frac{(1 - \lambda + \lambda^2)^3}{\lambda^2 (1-\lambda)^2},$$

so that the λ-line is a 6-fold cover of the j-line, with ramification at $j = 0$ (with index 3) and $j = 1728$ (with index 2).

1.4 Quotients of the upper half-plane as algebraic curves

More generally, suppose $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a subgroup of finite index. For simplicity let's assume that Γ acts on $\mathcal{H}$ without fixed points. Define the completed upper half-plane by setting $\mathcal{H}^* = \mathcal{H} \cup \mathbb{P}^1(\mathbb{Q})$. Then $\mathbf{P}^1(\mathbb{Q}) \subset \mathcal{H}^*$ is the set of *cusps*. $\mathcal{H}^*$ admits an action of $\mathrm{SL}_2(\mathbb{Q})$. It is given a topology by declaring a basis of neighborhoods of $\infty \in \mathcal{H}^*$ to be $\{\Im z > N\} \cup \{\infty\}$, for $N = 1, 2, \ldots$; bases around the other cusps are given by translation using $\mathrm{SL}_2(\mathbb{Z})$, which acts transitively on the set of cusps.

Proposition 1.4.1. *The complex structure on $\Gamma\backslash\mathcal{H}$ extends to a complex structure on $\Gamma\backslash\mathcal{H}^*$, which turns the latter into a compact Riemann surface.*

Proof. (Just a sketch.) We need to define the complex structure around the cusps. It suffices to do this for the cusp ∞. Let Γ_∞ be the stabilizer of ∞ in Γ. Then Γ_∞ takes the form

$$\Gamma_\infty = \begin{pmatrix} 1 & N\mathbb{Z} \\ 0 & 1 \end{pmatrix}$$

for some $N \geq 1$. The function $e^{2\pi i z/N}$ is a well-defined homeomorphism from a neighborhood of ∞ in $\Gamma\backslash\mathcal{H}$ to a neighborhood of $\mathbb{C}$, and this provides the complex structure.

$\Gamma\backslash\mathcal{H}^*$ is compact because it's a finite-to-one cover of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}^* = \mathbb{P}^1(\mathbb{C})$. □

Recall some basic facts about the relationship between compact Riemann surfaces, projective curves over $\mathbb{C}$, and complex function fields (meaning finite extensions of $\mathbb{C}(x)$):

- If X is a smooth projective curve over $\mathbb{C}$, then $X(\mathbb{C})$ has a natural structure of a compact Riemann surface.
- If S is a Riemann surface, then there is a smooth projective curve X over $\mathbb{C}$ with $X(\mathbb{C}) = S$, and X is unique up to isomorphism.
- If S is a Riemann surface whose corresponding to the smooth projective curve X, then the field of meromorphic functions on S is equal to the function field of X. In particular, the field of meromorphic functions on $\mathbb{P}^1(\mathbb{C})$ is $\mathbb{C}(x)$.

- The following categories are equivalent: compact Riemann surfaces, smooth projective curves over $\mathbb{C}$, and function fields over $\mathbb{C}$ (meaning finite extensions of $\mathbb{C}(x)$).

Thus there exists a complex projective curve $X(\Gamma)$ and a Zariski open subset $Y(\Gamma) \subset X(\Gamma)$ such that there is an isomorphism of complex manifolds $\Gamma\backslash\mathcal{H} \cong Y(\Gamma)(\mathbb{C})$.

2 Modular curves over number fields

We have just seen that for any finite-index subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$, the quotient $\Gamma\backslash\mathcal{H}^*$ is a compact Riemann surface and therefore corresponds to a smooth projective curve $X(\Gamma)$. In fact $X(\Gamma)$ always admits a model over a number field (exercise). For an important special class of Γ, however, $X(\Gamma)$ admits a model over $\mathbb{Q}$.

For $N \geq 1$, let $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$ be the group

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; c \equiv 0 \pmod{N} \right\}.$$

Theorem 2.0.2. *There exists a smooth projective curve $X_0(N)$ over $\mathbb{Q}$ whose complex points are $\Gamma_0(N)\backslash\mathcal{H}^*$.*

Proof. (Sketch) The functions $j(z)$ and $j(Nz)$ on $\mathcal{H}$ are both well-defined meromorphic functions on $\Gamma_0(N)\backslash\mathcal{H}^*$. Let S be a set of coset representatives for the quotient $\mathrm{SL}_2(\mathbb{Z})/\Gamma_0(N)$, and consider the polynomial

$$\prod_{\gamma \in S} (Y - j(N\gamma z)),$$

whose coefficients are a priori meromorphic functions on $\Gamma_0(N)\backslash\mathcal{H}^*$. But since these are symmetric functions in the $j(N\gamma z)$, they are well-defined on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}^* = \mathbb{P}^1$ and therefore they are rational functions in j. In fact, since each $j(N\gamma z)$ is holomorphic on $\mathcal{H}$, these coefficients must be polynomials in j. Therefore there is a bivariate polynomial $F(X, Y) \in \mathbb{C}[X, Y]$ such that

$$F(j(z), Y) = \prod_{\gamma \in S} (Y - j(N\gamma z)).$$

In particular, $F(j(z), j(Nz)) = 0$.

Then one shows that $F(X,Y)$ is irreducible and has integer coefficients (exercise). Let $X_0(N)$ be the smooth projective curve over $\mathbb{Q}$ whose function field is the fraction field of $\mathbb{Q}[X,Y]/F(X,Y)$. This means that some open affine $U \subset X_0(N)$ is isomorphic to the subvariety of $\mathbb{A}^2_{\mathbb{Q}}$ obtained by deleting the singular points from the plane curve $F(X,Y)=0$. We get a holomorphic map $\Gamma_0(N)\backslash\mathcal{H} \to U(\mathbb{C})$ defined by $z \mapsto (j(z), j(Nz))$. This map is injective (exercise). It extends to a map of compact Riemann surfaces $\Gamma_0(N)\backslash\mathcal{H}^* \to X_0(N)(\mathbb{C})$, which (since the map is nonconstant and injective and $X_0(N)_{\mathbb{C}}$ is irreducible) has to be an isomorphism. $\square$

We remark that if $\Gamma(N)$ is the principal congruence subgroup

$$\Gamma(N) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\},$$

then $\Gamma(N)\backslash\mathcal{H}^*$ has a model over the cyclotomic field $\mathbb{Q}(\zeta_N)$.