Modular curves and Galois representations

Notes for a mini-course held at the workshop *Modular Forms and Related Topics II*
Center for Advanced Mathematical Sciences, American University of Beirut
6–10 February 2012

Peter Bruin (Universität Zürich)
peter.bruin@math.uzh.ch

Abstract: The first half of the mini-course is an introduction to various aspects of modular curves and modular forms. We will relate the classical description of modular curves (as quotients of the complex upper half-plane) to moduli of elliptic curves. This leads to a 'finer' description of modular curves as algebraic curves over the rationals. In the second half of the mini-course, we will describe how modular curves and their Jacobians can be used to attach two-dimensional Galois representations to modular forms, in particular over finite fields. We will finish with some words on how all of this can be made computable.

Introduction

The goal of these notes (a written and slightly expanded version of the lectures given at CAMS) is to give an overview of a few aspects of modular curves and of Galois representations attached to modular forms.

In the first two talks, we give a brief (and necessarily very incomplete) introduction to elliptic curves and modular curves from both an analytic and an algebraic perspective. We start by defining modular curves in the classical way as quotients of the complex upper half-plane, and we relate this description to moduli of complex tori. Next, we explain how complex tori can be viewed in an algebraic way as elliptic curves, and we extend the definition of elliptic curves to arbitrary fields. This leads to a definition of modular curves as algebraic curves over $\mathbf{Q}$ classifying elliptic curves with extra structure.

This 'arithmetic' structure of modular curves can be used as the starting point for associating two-dimensional representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ to modular forms. We refer to Gabor Wiese's lectures for a general introduction to Galois representations. In the third talk, we introduce Galois representations attached to modular forms. In the fourth talk, we focus on modular forms over finite fields. We describe how the associated Galois representations appear in the Jacobian varieties of modular curves. We finish by explaining the results of Edixhoven, Couveignes [13] and the author [1] on computing modular Galois representations over finite fields and coefficients of modular forms over number fields.

There exists an enormous amount of mathematical literature on the subject of modular curves, modular forms and Galois representations. Let us just mention a few recommended references. An accessible and extensive reference treating the material in the first half of the mini-course, and much more, is the book of Diamond and Shurman [11]. For the analytic theory, the book of Miyake [20] is also recommended. Slightly more advanced references are the survey [10] of Diamond and Im, which is extremely useful and has an extensive bibliography, and the book of Ribet and Stein [22]. For anybody who is seriously interested in modular forms, Shimura's influential book [24] is impossible to ignore. More advanced treatments of modular curves from an algebraic point of view can be found in Deligne and Rapoport [8], Katz and Mazur [15], and Conrad [5]. The topic of elliptic curves, essential for a good understanding of modular curves, is even broader. The books of Silverman [25], [26], Silverman and Tate [27], Cassels [4] and Milne [19], as well as the article of Tate [28], are just a few of the many introductions. The construction of Galois representations attached to modular forms can be found for modular forms of weight at least 2 in Deligne's article [7] and for weight 1 in Deligne and Serre [9]. A more expanded description of Deligne's construction will be given in the future in a book of Conrad [6].

Acknowledgements. I would like to thank Kamal Khuri-Makdisi, Wissam Raji and the staff at CAMS for their efforts in organising this workshop and for their hospitality. I also thank all the participants of the workshop for a very enjoyable week in Lebanon.

1. Modular curves: complex analytic aspects

References: Diamond and Im [10, § 7] and the references therein; Diamond and Shurman [11, Chapter 1].

In this first talk, we start by defining modular curves as quotients of the complex upper half-plane by a certain kind of group action. We then give a "moduli interpretation" of modular curves as geometric objects classifying other geometric objects (namely, complex tori).

1.1. Modular curves as Riemann surfaces

The two fundamental objects in the theory of modular curves and modular forms are the complex upper half-plane

$$\mathfrak{H} = \{\tau \in \mathbf{C} \mid \Im\tau > 0\}$$

and the group

$$\mathrm{SL}_2(\mathbf{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, a, b, c, d \in \mathbf{Z}, ad - bc = 1 \right\}.$$

They are connected by a left action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathfrak{H}$ defined by

$$\gamma\tau = \frac{a\tau + b}{c\tau + d} \quad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}). \tag{1.1}$$

For every positive integer n, we define a subgroup $\Gamma(n) \subseteq \mathrm{SL}_2(\mathbf{Z})$ by

$$\Gamma(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) \middle| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod n \right\}.$$

Definition. A *congruence subgroup* is a subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbf{Z})$ that contains $\Gamma(n)$ for some n.

The most important congruence subgroups are $\Gamma_0(n)$ and $\Gamma_1(n)$, consisting of those matrices in $\mathrm{SL}_2(\mathbf{Z})$ whose reduction modulo n are of the form $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$, respectively, where $*$ stands for any element of $\mathbf{Z}/n\mathbf{Z}$.

Remark. Every congruence subgroup has finite index in $\mathrm{SL}_2(\mathbf{Z})$, but the converse does not hold.

Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$. We define

$$\mathrm{Y}_\Gamma = \Gamma\backslash\mathfrak{H}.$$

This is a non-compact Riemann surface (one has to be careful at the points of $\mathfrak{H}$ that have non-trivial stabiliser in Γ).

1.2. Lattices and complex tori

Definition. A *lattice* (in $\mathbf{C}$) is a subgroup $L \subset \mathbf{C}$ generated by two elements $z_1, z_2 \in \mathbf{C}$ that are linearly independent over $\mathbf{R}$.

We note that such an L is isomorphic to $\mathbf{Z}^2$ as an Abelian group. We do not view the points z_1 and z_2 as part of the data defining L, so there is no distinguished identification of L with $\mathbf{Z}^2$.

Definition. Two lattices L and L' are *homothetic* if there exists $\lambda \in \mathbf{C}^\times$ such that $L' = \lambda \cdot L$.

Let L be a lattice. The quotient $\mathbf{C}/L$ has the structure of an Abelian group as well as the structure of a Riemann surface. We will denote the image of a point $z \in \mathbf{C}$ under the quotient map $\mathbf{C} \to \mathbf{C}/L$ by $[z]$.

Fact 1.1. *Let L and L' be lattices in $\mathbf{C}$, and let $h\colon \mathbf{C}/L \to \mathbf{C}/L'$ be a holomorphic map sending $[0]$ to $[0]$. Then there exists $\lambda \in \mathbf{C}$ with $\lambda \cdot L \subseteq L'$ such that the diagram*

$$\begin{tikzcd} \mathbf{C} \arrow[r, "\lambda\cdot"] \arrow[d] & \mathbf{C} \arrow[d] \\ \mathbf{C}/L \arrow[r, "h"] & \mathbf{C}/L' \end{tikzcd}$$

is commutative.

One can use this to show that every holomorphic map $\mathbf{C}/L \to \mathbf{C}/L'$ sending $[0]$ to $[0]$ is a group homomorphism, and that the group $\mathrm{Hom}(\mathbf{C}/L, \mathbf{C}'/L)$ of all such maps is canonically isomorphic to the group $\{\lambda \in \mathbf{C} \mid \lambda L \subseteq L'\}$.

Taking $L = L'$, we see that the holomorphic maps $\mathbf{C}/L \to \mathbf{C}/L$ preserving $[0]$ form a ring that is isomorphic to $\{\lambda \in \mathbf{C} \mid \lambda L \subseteq L\}$.

Definition. A *complex torus* is a Riemann surface T together with a distinguished point $O \in T$ such that there exist a lattice $L \subset \mathbf{C}$ and an isomorphism of Riemann surfaces (holomorphic map admitting a holomorphic inverse)

$$\phi\colon \mathbf{C}/L \xrightarrow{\sim} T$$

that sends $[0]$ to O.

Remark. Equivalently, a complex torus is a compact connected Riemann surface T of genus 1 together with a point $O \in T$.

Any isomorphism ϕ as above gives T the structure of an Abelian group. Using Fact 1.1, one can show that this structure does not depend on the choice of ϕ. Given (T, O), the lattice L is unique up to homothety.

1.3. Moduli interpretation of the upper half-plane

We consider group homomorphisms

$$\omega\colon \mathbf{Z}^2 \rightarrowtail \mathbf{C}$$

whose image in $\mathbf{C}$ is a lattice. Given such an ω, we abbreviate

$$\omega_1 = \omega\begin{pmatrix}1\\0\end{pmatrix}, \quad \omega_2 = \omega\begin{pmatrix}0\\1\end{pmatrix}.$$

We write

$$\begin{aligned}\mathfrak{M} &= \bigl\{\omega\colon \mathbf{Z}^2 \rightarrowtail \mathbf{C} \bigm| \Im(\omega_1/\omega_2) > 0\bigr\} \\ &\cong \{(z_1, z_2) \in \mathbf{C}^\times \times \mathbf{C}^\times \mid \Im(z_1/z_2) > 0\}.\end{aligned}$$

This can be thought of as the space of lattices in $\mathbf{C}$ together with a "negatively oriented" basis. The second description gives $\mathfrak{M}$ the structure of a complex manifold of (complex) dimension 2. There is an action of $\mathbf{C}^\times$ on $\mathfrak{M}$ given by

$$(\lambda\omega)(v) = \lambda \cdot \omega(v).$$

There is a surjective map

$$\begin{aligned}Q\colon \mathfrak{M} &\to \mathfrak{H} \\ \omega &\mapsto \omega_1/\omega_2.\end{aligned}$$

This is a quotient map for the action of $\mathbf{C}^\times$ on $\mathfrak{M}$; in other words, two elements $\omega, \omega' \in \mathfrak{M}$ have the same image under Q if and only if there exists $\lambda \in \mathbf{C}^\times$ such that $\omega' = \lambda \cdot \omega$.

Remark. One can also think of $\mathfrak{M}$ in various other ways:

(1) as the space of negatively oriented bases of $\mathbf{C}$ as an $\mathbf{R}$-vector space;

(2) as the space of isomorphism classes of triples (V, α, β) with V a one-dimensional complex vector space, $\alpha\colon \mathbf{C} \xrightarrow{\sim} V$ an isomorphism, and $\beta\colon \mathbf{Z}^2 \rightarrowtail V$ a lattice together with a negatively oriented basis. In this interpretation, the map $Q\colon \mathfrak{M} \to \mathfrak{H}$ means forgetting the isomorphism α.

(3) as the group $\mathrm{GL}_2(\mathbf{R})^+$ of 2×2-matrices with real coefficients and positive determinant, via the isomorphism of real manifolds

$$\begin{aligned}\mathrm{GL}_2(\mathbf{R})^+ &\xrightarrow{\sim} \mathfrak{M} \\ \begin{pmatrix}a & b\\ c & d\end{pmatrix} &\longmapsto (\omega_1 = ai + b, \omega_2 = ci + d).\end{aligned}$$

1.4. The action of $\mathrm{SL}_2(\mathbf{Z})$

We recall that the group $\mathrm{SL}_2(\mathbf{Z})$ acts from the left on $\mathfrak{H}$ by (1.1). On the other hand, from the standard action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathbf{Z}^2$ we obtain a left action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathfrak{M}$ defined by

$$\gamma\omega = \omega \circ \gamma^{\mathrm{t}} \quad \text{for all } \gamma \in \mathrm{SL}_2(\mathbf{Z}),\ \omega \in \mathfrak{M},$$

where γ^{t} is the transpose of γ. When $\gamma = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$, this means

$$(\gamma\omega)_1 = a\omega_1 + b\omega_2, \quad (\gamma\omega)_2 = c\omega_1 + d\omega_2.$$

We note that the actions of $\mathbf{C}^\times$ and $\mathrm{SL}_2(\mathbf{Z})$ on $\mathfrak{M}$ are compatible with each other.

Lemma 1.2. *The quotient map $Q\colon \mathfrak{M} \to \mathfrak{H}$ is $\mathrm{SL}_2(\mathbf{Z})$-equivariant.*

Proof. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$, and let $\omega \in \mathfrak{M}$. We compute

$$Q(\gamma\omega) = \frac{(\gamma\omega)_1}{(\gamma\omega)_2} = \frac{a\omega_1 + b\omega_2}{c\omega_1 + d\omega_2}$$

and

$$\gamma Q(\omega) = \frac{aQ(\omega) + b}{cQ(\omega) + d} = \frac{a\omega_1/\omega_2 + b}{c\omega_1/\omega_2 + d}.$$

From the equality of these two expressions, we conclude that Q is compatible with the action of $\mathrm{SL}_2(\mathbf{Z})$. □

1.5. Moduli interpretation of modular curves

For concreteness, we now restrict ourselves to congruence subgroups of the form $\Gamma_1(n)$. If in addition n is at least 4, then $\Gamma_1(n)$ acts without fixed points on $\mathfrak{H}$, but we will not use this. We write

$$\mathrm{Y}_1(n) = \mathrm{Y}_{\Gamma_1(n)} = \Gamma_1(n)\backslash\mathfrak{H}.$$

We will shortly give the moduli interpretation of $\mathrm{Y}_1(n)$, but we start with $\Gamma_1(n)\backslash\mathfrak{M}$.

For all $\omega \in \mathfrak{M}$, we write L_ω for the image of ω viewed as a map $\mathbf{Z}^2 \rightarrowtail \mathbf{C}$, and we write

$$P_\omega = [\omega_2/n] \in \mathbf{C}/L_\omega.$$

Theorem 1.3. *There is a bijection*

$$\begin{aligned} \Gamma_1(n)\backslash\mathfrak{M} &\xrightarrow{\sim} \{(L, P) \mid L \subset \mathbf{C} \textit{ lattice}, P \in \mathbf{C}/L \textit{ of order } n\} \\ \Gamma_1(n)\omega &\longmapsto (L_\omega, P_\omega). \end{aligned} \tag{1.2}$$

Proof. We start by noting that the pair (L_ω, P_ω) is indeed an element of the right-hand side of (1.2). One can check (although this is not completely trivial) that every element on the right-hand side can be obtained in this way. The proof of this fact is left as an exercise.

It remains to see under which conditions two elements $\omega, \omega' \in \mathfrak{M}$ give the same element of the right-hand side of (1.2). First, the lattices L_ω and $L_{\omega'}$ are equal if and only if there exists $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ such that $\omega' = \gamma\omega$, where $\gamma\omega = \omega \circ \gamma^{\mathrm{t}}$ as before. Such a γ, if it exists, is uniquely determined by ω and ω'. Thus in order to have $(L_\omega, P_\omega) = (L_{\omega'}, P_{\omega'})$, the existence of such a γ is a necessary condition. Given such a $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have the following equivalences:

$$\begin{aligned} P_{\omega'} = P_\omega &\iff (\gamma\omega)_2/n \equiv \omega_2/n \pmod{L_\omega} \\ &\iff (\gamma\omega)_2 \equiv \omega_2 \pmod{nL_\omega} \\ &\iff \omega\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \omega\begin{pmatrix} 0 \\ 1 \end{pmatrix} \pmod{nL_\omega} \\ &\iff \begin{pmatrix} c \\ d \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \pmod{n\mathbf{Z}^2} \\ &\iff \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{n}. \end{aligned}$$

This proves the theorem. □

Generalising the concept of homothety of lattices, we say that two such pairs (L, P) and (L', P') are *homothetic*, notation $(L, P) \sim (L', P')$, if there exists $\lambda \in \mathbf{C}^\times$ such that $L' = \lambda L$ and $P' = \lambda P$. (Note that λP is a well-defined element of $\mathbf{C}/L'$ because $L' = \lambda L$.)

Theorem 1.4. *There is a bijection*

$$\begin{aligned} \mathrm{Y}_1(n) &\xrightarrow{\sim} \{(L, P) \mid L \subset \mathbf{C} \textit{ lattice}, P \in \mathbf{C}/L \textit{ of order } n\}/\!\sim \\ \Gamma_1(n)\tau &\longmapsto [\mathbf{Z}\tau + \mathbf{Z}, [1/n]]. \end{aligned}$$

Proof. This follows from Theorem 1.3 by taking the quotient by the action of $\mathbf{C}^\times$ on $\Gamma_1(n)\backslash\mathfrak{M}$, and by the relation of homothety on the set of pairs (L, P). □

Corollary 1.5. *There is a bijection*

$$\begin{aligned} \mathrm{Y}_1(n) &\xrightarrow{\sim} \{\textit{complex tori with a point of order } n\}/\cong \\ \Gamma_1(n)\tau &\longmapsto [\mathbf{C}/(\mathbf{Z}\tau + \mathbf{Z}), [1/n]]. \end{aligned}$$

Proof. This follows from the fact that homothety classes of pairs (L, P) as above are the same as isomorphism classes of complex tori with a point of order n. □

1.6. Compactifications

Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$. The non-compact Riemann surface $\mathrm{Y}_\Gamma = \Gamma\backslash\mathfrak{H}$ can made into a compact Riemann surface by adding a finite number of points, called *cusps*. This compactification is denoted by X_Γ.

There also exists a moduli interpretation of the cusps, related to 'degenerating' lattices and tori. We will not go into the details.

2. Modular curves: algebraic and arithmetic aspects

References: Diamond and Im [10, §§ 8–9], and the references therein; Diamond and Shurman [11, Chapter 7].

2.1. Elliptic curves

Definition. Let L be a lattice in $\mathbf{C}$. The *Weierstrass $\wp$-function* associated to L is the meromorphic function

$$\wp_L\colon \mathbf{C} \to \mathbf{P}^1(\mathbf{C}) = \mathbf{C} \cup \{\infty\}$$

defined for $z \neq L$ by the infinite sum

$$\wp_L(z) = \frac{1}{z^2} + \sum_{\substack{\omega\in L\\ \omega\neq 0}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right)$$

and for $z \in L$ by $\wp_L(z) = \infty$. The sum converges absolutely and uniformly on compact subsets of $\mathbf{C} - L$. The function $\wp_L$ has poles of order 2 at the points of L.

Lemma 2.1. *The function $\wp_L$ is even and invariant under translation by elements of L, i.e. for all $z \in \mathbf{C}$ and $\omega \in L$ it satisfies*

$$\wp_L(-z) = \wp_L(z)$$

and

$$\wp_L(z + \omega) = \wp_L(z).$$

Proof. The claim that $\wp_L$ is even easily follows from the definition by using te fact that the map $\omega \mapsto -\omega$ is a bijection from L to itself. To prove the claim that $\wp_L$ is invariant under translation, we first compute the derivative of $\wp_L$ as

$$\wp_L'(z) = -2\sum_{\omega\in L} \frac{1}{(z-\omega)^3}.$$

This is clearly invariant under translation by elements of L. We deduce that for every $\omega \in L$ there is $c_\omega \in L$ such that

$$\wp_L(z + \omega) = \wp(z) + c_L \quad \text{for all } z \in \mathbf{C}.$$

Putting $z = -\omega/2$ and comparing with the identity $\wp_L(-\omega/2) = \wp_L(\omega/2)$, we conclude that $c_L = 0$. □

Theorem 2.2. *The functions $\wp_L$ and $\wp'_L$ satisfy an equation of the form*

$$\wp'_L(z)^2 = 4\wp_L(z)^3 - g_2\wp_L(z) - g_3,$$

where g_2 and g_3 are certain complex numbers depending on L.

Proof. We expand each term in the definition of $\wp_L$ in a Laurent series in z, change the order of summation, and take a suitable linear combination of the resulting series for $\wp_L(z)^3$, $\wp_L(z)$ and $\wp'_L(z)^2$ to make the non-positive powers of z cancel. (A priori, we would also need $\wp_L(z)^2$, but this turns out to be unnecessary.) For suitable g_2 and g_3, we obtain

$$\wp'_L(z)^2 - 4\wp_L(z)^3 + g_2\wp_L(z) + g_3 = O(z^2) \quad \text{as } z \to 0.$$

Now the left-hand side can be extended to a holomorphic L-invariant function h on all of $\mathbf{C}$. This h is bounded, because it assumes all its values already on the closure of a fundamental parallellogram, which is compact. By Liouville's theorem, h is constant, and substituting $z = 0$ shows that $h = 0$. It turns out that the correct values for g_2 and g_3 are

$$g_2 = 60G_4(L) \quad \text{and} \quad g_3 = 140G_6(L),$$

where

$$G_k(L) = \sum_{\substack{\omega\in L\\ \omega\neq 0}} \frac{1}{\omega^k}.$$

The details are left as an exercise. □

Remark. For $L = \mathbf{Z}\tau + \mathbf{Z}$, the function $G_k(L)$ equals, up to multiplication by a constant, the Eisenstein series $E_k(\tau)$ of weight k defined in Prof. Kohnen's first lecture.

Via the functions $\wp_L$ and $\wp'_L$, the complex torus $\mathbf{C}/L$ can be given the structure of an algebraic curve in the projective plane $\mathbf{P}^2(\mathbf{C})$. More precisely, we have an embedding of complex manifolds

$$\begin{aligned} \psi_L\colon \mathbf{C}/L &\rightarrowtail \mathbf{P}^2(\mathbf{C}) \\ [z] &\mapsto \begin{cases} (\wp_L(z) : \wp'_L(z) : 1) & \text{if } z \notin L; \\ (0 : 1 : 0) & \text{if } z \in L. \end{cases} \end{aligned}$$

The image is the complex curve (or Riemann surface) E_L given by the homogeneous cubic equation

$$Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3.$$

We conclude that from L we have constructed a cubic curve E_L in $\mathbf{P}^2(\mathbf{C})$ together with a distinguished point, namely $(0 : 1 : 0)$. Via the map ψ_L, we can identify $\mathbf{C}/L$ with E_L. Because $\mathbf{C}/L$ is an Abelian group with neutral element $[0]$, E_L also gets a group structure with neutral element $(0 : 1 : 0)$. We call E_L the *elliptic curve* associated to L.

Remark. Usually, one uses the affine equation because it is shorter. That is to say, one simply writes

$$E\colon y^2 = 4x^3 - g_2x - g_3.$$

We now consider *cubic curves* over an arbitrary field K, given by a homogeneous equation

$$C\colon \sum_{\substack{i,j,k\geq 0\\ i+j+k=3}} c_{i,j,k}X^iY^jZ^k = 0 \quad \text{with } c_{i,j,k} \in K. \tag{2.1}$$

It does not matter very much whether we view C as a set of points in $\mathbf{P}^2(\bar{K})$ (where $\bar{K}$ is an algebraic closure of K), as a function field over K, or as a scheme over K.

We say that C is *smooth* if the three partial derivatives do not vanish simultaneously anywhere on the curve. We will denote by $C(K)$ the set of K-rational points of C. This is the set of solutions of the homogeneous equation (2.1) in $\mathbf{P}^2(K)$.

Definition. Let K be a field. An *elliptic curve* over K is a smooth cubic curve E over K together with a point $O \in E(K)$.

Fact 2.3. *Up to a change of coordinates, every elliptic curve E is given by an equation of the form*

$$E\colon Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3, \tag{2.2}$$

where a_1, a_2, a_3, a_4, a_6 are elements of K and O corresponds to the point $(0:1:0)$.

An equation of the form (2.2) is called a *Weierstrass equation*. Usually, E is denoted instead by the affine equation

$$E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{2.3}$$

The smoothness of E is equivalent to the non-vanishing of the *discriminant* of the above equation, which is a certain polynomial in the coefficients $a_1, \ldots, a_6$. The set $E(K)$ of K-rational points of E now consists of the solutions of the homogeneous equation (2.2) in $\mathbf{P}^2(K)$. These correspond bijectively to the solutions of (2.3) in $K \times K$ together with the point at infinity.

Remark. If the characteristic of K is not 2 or 3, one can make a change of variables giving a simpler equation

$$E\colon y^2 = x^3 + px + q.$$

In this case the smoothness of E is equivalent to the condition $4p^3 + 27q^2 \neq 0$.

Any cubic curve has the property that every line intersects the curve in exactly three points, counted with multiplicity. Let E be an elliptic curve over any field, embedded into the projective plane by a Weierstrass equation. Then the group structure on E is characterised by the property that three points of E add up to 0 if and only if they lie on a line.

Let (E, O) and (E', O') be elliptic curves. One can show that every map of algebraic curves $E \to E'$ sending O to O' preserves the group structure; this is analogous to the similar statement about holomorphic maps between complex tori.

2.2. Algebraic modular curves

In the previous lecture, we saw that modular curves over $\mathbf{C}$ parametrise complex tori with "level structure". We have just seen that complex tori can be viewed as elliptic curves over $\mathbf{C}$. This interpretation allows us to give an algebraic interpretation of modular curves, and to define them over $\mathbf{Q}$, or more generally over $\mathbf{Z}[1/n]$ for some n.

An (algebraic) modular curve is a curve which classifies elliptic curves over fields with additional level structure. For concreteness, we restrict to the case where the level structure consists of a rational point of order n, where $n \geq 4$.

Fact 2.4 (somewhat imprecise). *Let $n \geq 4$. There exists a smooth affine curve $\mathrm{Y}_1(n)$ defined over $\mathbf{Z}[1/n]$ such that for any field K whose characteristic does not divide n, the set $\mathrm{Y}_1(n)(K)$ of K-rational points of $\mathrm{Y}_1(n)$ is in bijection with the set of isomorphism classes of pairs (E, P) consisting of an elliptic curve E over K and a point $P \in E(K)$ of order n.*

Remark. Using the language of schemes, one can define $\mathrm{Y}_1(n)$, together with a so-called universal elliptic curve over it, as a universal object in the category of elliptic curves over schemes together with a section which is everywhere of order n.

Example. We consider elliptic curves with a point of order 4 over a field K of characteristic different from 2. Consider an elliptic curve E over K and a point $P \in E(K)$ of order 4. We note that $2P$ is a point of order 2. After a suitable choice of coordinates, we may assume that E is given by an equation

$$E\colon y^2 = x(x^2 + ax + b)$$

and that $2P$ has coordinates $(0, 0)$. Let (r, s) be the coordinates of P, and consider the invariant

$$\mu(E, P) = r^3/s^2 \in K.$$

One can show that this is independent of the choice of coordinates. We get a bijection

$$\{\text{elliptic curves with a point of order 4 over } K\}/{\cong} \xrightarrow{\sim} K - \{0, 1/4\}$$

mapping (E, P) to $\mu(E, P)$. The inverse is given by

$$\mu \mapsto (E\colon y^2 = x(x^2 + (1 - 2c)x + c^2), P = (c, c)).$$

This shows that the modular curve $\mathrm{Y}_1(n)$ over $\mathbf{Z}[1/n]$ parametrising elliptic curves with a point of order 4 is the affine line with two points removed, or equivalently the projective line with three points removed.

2.3. Hecke correspondences

Let $n \geq 4$ be an integer. For any positive integer m, let $\mathrm{Y}_1(n; m)$ denote the modular curve classifying triples (E, P, C), where E is an elliptic curve, P is a point of order n, and C is a (not necessarily cyclic) subgroup of order m such that $\langle P \rangle \cap C = \{0\}$. In terms of congruence subgroups, $\mathrm{Y}_1(n; m)$ is defined by the group

$$\Gamma_1(n; m) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(n) \,\middle|\, b \equiv 0 \pmod{m} \right\}.$$

The m-th *Hecke correspondence* on $\mathrm{Y}_1(n)$ is the diagram

$$\begin{tikzcd} & \mathrm{Y}_1(n; m) \arrow[dl, "q_1"'] \arrow[dr, "q_2"] & \\ \mathrm{Y}_1(n) & & \mathrm{Y}_1(n) \end{tikzcd} \tag{2.4}$$

where q_1 and q_2 send the point of $\mathrm{Y}_1(n; m)$ corresponding to a triple (E, P, C) to the points of $\mathrm{Y}_1(n)$ corresponding to the pairs (E, P) and $(E/C, P \bmod C)$, respectively.

For all $d \in (\mathbf{Z}/n\mathbf{Z})^\times$, the d-th *diamond automorphism* is the automorphism

$$r_d\colon \mathrm{Y}_1(n) \to \mathrm{Y}_1(n) \tag{2.5}$$

sending the point corresponding to a pair (E, P) to the point corresponding to the pair (E, dP).

The Hecke correspondences and diamond automorphisms can be used to define endomorphisms (called *Hecke operators* and *diamond operators*, respectively) on spaces of modular forms for $\Gamma_1(n)$, as well as on (co)homology groups and on the Jacobian of $\mathrm{Y}_1(n)$.

2.4. Compactified modular curves

Let $n \geq 4$ be an integer. The modular curve $\mathrm{Y}_1(n)$ is an affine curve. Up to isomorphism, there is a unique smooth projective curve over $\mathbf{Z}[1/n]$ containing $\mathrm{Y}_1(n)$ as an open subset. This compactification is denoted by $\mathrm{X}_1(n)$. The set $\mathrm{X}_1(n)(\mathbf{C})$ of complex points of $\mathrm{X}_1(n)$ is isomorphic to the compactification of $\Gamma_1(n)\backslash\mathfrak{H}$ that we saw earlier.

3. Galois representations in Jacobians of modular curves

References: Mazur [18], Ribet [21], Gross [14], Edixhoven [12].

3.1. Hecke algebras and eigenforms

Let n and k be positive integers. Let $\mathrm{M}_k(n)$ denote the space of modular forms of weight k for the group $\Gamma_1(n)$, and let $\mathbf{T}(\mathrm{M}_k(n))$ denote the Hecke algebra acting on $\mathrm{M}_k(n)$. We recall that $\mathbf{T}(\mathrm{M}_k(n))$ is a commutative ring that is generated as a $\mathbf{Z}$-algebra by the *Hecke operators* T_m for all $m \geq 1$, and that $\mathbf{T}(\mathrm{M}_k(n))$ is free of finite rank as a $\mathbf{Z}$-module.

Definition. Let K be a field. A *(Hecke) eigenform* of weight k for the group $\Gamma_1(n)$ with coefficients in k is an element $f \in \mathrm{M}_k(n)(K)$ that is an eigenvector for all the Hecke operators T_m with $m \geq 1$.

Remark. The ring $\mathbf{T}(\mathrm{M}_k(n))$ is also generated by the T_p for all prime numbers p and the *diamond operators* $\langle d \rangle$ for all $d \in (\mathbf{Z}/n\mathbf{Z})^\times$. If f is an eigenform, then it has a well-defined *character* (sometimes called *Nebentypus*), which is a group homomorphism

$$\epsilon\colon (\mathbf{Z}/n\mathbf{Z})^\times \to \mathbf{C}^\times$$

determined uniquely by the property that f is an eigenvector of $\langle d \rangle$ with eigenvalue $\epsilon(d)$ for all $d \in (\mathbf{Z}/n\mathbf{Z})^\times$.

To every eigenform $f \in \mathrm{M}_k(n)(K)$, we associate a ring homomorphism

$$\phi_f\colon \mathbf{T}(\mathrm{M}_k(n)) \to K$$

mapping every operator to its eigenvalue on f. If f is normalised such that its first coefficient $a_1(f)$ equals 1, then $\phi_f(T_m) = a_m(f)$ for all $m \geq 1$.

3.2. Galois representations attached to eigenforms

Theorem 3.1 (Eichler, Shimura, Igusa, Deligne, Serre). *Let $f \in \mathrm{M}_k(n)$ be a normalised Hecke eigenform with character ϵ. Let $\mathbf{Q}_f = \mathbf{Q}(\{a_m(f) \mid m \ge 1\})$ denote the number field generated by the coefficients of f, and consider ϵ as a group homomorphism $(\mathbf{Z}/n\mathbf{Z})^\times \to \mathbf{Q}_f^\times$. Let λ be a finite place of $\mathbf{Q}_f$ of residue characteristic l, and let $\mathbf{Q}_{f,\lambda}$ denote the completion of $\mathbf{Q}_f$ with respect to λ. There exists a continuous semi-simple group homomorphism*

$$\rho_{f,\lambda}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$$

with the following properties:

(1) $\rho_{f,\lambda}$ is unramified outside nl;

(2) for every prime number $p \nmid nl$, the characteristic polynomial of $\rho_{f,\lambda}(\sigma_p)$, where σ_p is a Frobenius element at p, equals $t^2 - a_p(f)t + \epsilon(p)p^{k-1}$.

Furthermore, this $\rho_{f,\lambda}$ is odd and unique up to conjugation in $\mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$.

Let us make some remarks to explain the conditions and conclusion of the theorem.

1. The group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is a compact topological group, where the closed subgroups correspond bijectively to the subfields of $\overline{\mathbf{Q}}$ via the *Galois correspondence*. This correspondence is obtained as follows: to every subfield K of $\overline{\mathbf{Q}}$ we associate the stabiliser

$$G_K = \{\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \mid \sigma(x) = x \text{ for all } x \in K\}$$

and to every closed subgroup H of G we associate the field of invariants

$$\overline{\mathbf{Q}}^H = \{x \in \overline{\mathbf{Q}} \mid \sigma(x) = x \text{ for all } \sigma \in H\}.$$

2. The completion $\mathbf{Q}_{f,\lambda}$ is a topological field; it is a finite extension of the field $\mathbf{Q}_l$ of l-adic numbers. The topology on $\mathbf{Q}_{f,\lambda}$ makes $\mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$ into a topologial group.

3. The kernel of $\rho_{f,\lambda}$ is the inverse image of a closed normal subgroup under a continuous homomorphism, and hence is a closed normal subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. The subfield $K_{f,\lambda}$ of $\overline{\mathbf{Q}}$ corresponding to $\ker\rho_{f,\lambda}$ is a Galois extension of $\mathbf{Q}$, and $\rho_{f,\lambda}$ factors as

$$\rho_{f,\lambda}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \twoheadrightarrow \mathrm{Gal}(K_{f,\lambda}/\mathbf{Q}) \rightarrowtail \mathrm{GL}_2(\mathbf{Q}_{f,\lambda}).$$

4. Let p be a prime number not dividing nl. One way of stating condition (1) is to say that the extension $K_{f,\lambda}/\mathbf{Q}$ is unramified at p; this means that the element $p \in K_{f,\lambda}$ has valuation 1 with respect to every extension of the p-adic valuation from $\mathbf{Q}$ to $K_{f,\lambda}$. An equivalent way of stating condition (1) is as follows. After choosing an embedding

$$i_p\colon \overline{\mathbf{Q}} \to \overline{\mathbf{Q}}_p,$$

we obtain an injection of Galois groups

$$r_p\colon \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p) \rightarrowtail \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}).$$

Via r_p, we identify $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ with a subgroup G_p of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Furthermore, there is a canonical homomorphism

$$s_p\colon G_p \to \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p).$$

The kernel of s_p is called the *inertia subgroup* of G_p and denoted by I_p. Now condition (1) means that $\rho_{f,\lambda}$ is trivial when restricted to I_p.

5. Recall that the *Frobenius automorphism* $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbf{F}}_p/\overline{\mathbf{F}})$ is the automorphism of $\overline{\mathbf{F}}_p$ defined by $\mathrm{Frob}_p(x) = x^p$ for all $x \in \overline{\mathbf{F}}_p$. Choose an element

$$\widetilde{\mathrm{Frob}_p} \in G_p \quad \text{such that} \quad s_p(\widetilde{\mathrm{Frob}_p}) = \mathrm{Frob}_p\,.$$

By condition (1), the matrix $\rho_{f,\lambda}(\widetilde{\mathrm{Frob}_p}) \in \mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$ is independent of the choice of $\widetilde{\mathrm{Frob}_p}$. We denoted this matrix by $\rho_{f,\lambda}(\mathrm{Frob}_p)$. This still depends on the choice of the embedding i_p: different choices of i_p lead to $\rho_{f,\lambda}(\mathrm{Frob}_p)$ that are conjugate in $\mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$.

6. Because $\rho_{f,\lambda}$ is well-defined up to conjugation, its characteristic polynomial is well-defined independently of any choices made. Condition (2) says that this characteristic polynomial must be equal to $t^2 - a_p(f)t + \epsilon(p)p^{k-1}$. Note that this polynomial actually lies in the subring $\mathbf{Q}_f[t]$ of $\mathbf{Q}_{f,\lambda}[t]$, has algebraically integral coefficients, and is independent of λ.

We recall that $\mathrm{M}_k(n)$ is the direct sum of the subspace $\mathcal{E}_k(n)$ of Eisenstein series and the subspace $\mathrm{S}_k(n)$ of cusp forms. Moreover, every eigenform is either an Eisenstein series or a cusp form. The fact that there are elementary formulae for the coefficients of Eisenstein series (e.g. $a_m(E_k) = \sum_{d|m} d^{k-1}$ for the Eisenstein series E_k of weight k for $\mathrm{SL}_2(\mathbf{Z})$) is reflected in the following fact about Galois representations.

Theorem 3.2. *Let $f \in \mathrm{M}_k(n)$ be a normalised eigenform, and let λ be a finite place of $\mathbf{Q}_f$.*

(1) If f is an Eisenstein series, then ρ_f is reducible, i.e. there are two continuous homomorphisms

$$\chi_1, \chi_2\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_1(\mathbf{Q}_{f,\lambda}) = \mathbf{Q}_{f,\lambda}^\times$$

such that $\rho_{f,\lambda}$ is described up to conjugation by a fixed matrix in $\mathrm{GL}_2(\mathbf{Q}_{f,\lambda})$ by

$$\rho_{f,\lambda}(\sigma) = \begin{pmatrix} \chi_1(\sigma) & 0 \\ 0 & \chi_2(\sigma) \end{pmatrix} \quad \text{for all } \sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}).$$

(2) If f is a cusp form, then $\rho_{f,\lambda}$ is irreducible, i.e. there is no vector in $\mathbf{Q}_{f,\lambda}^2 - \{0\}$ that is a common eigenvector of the $\rho_{f,\lambda}(\sigma)$ for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

Example. Let f be the Eisenstein series E_k of weight k for $\mathrm{SL}_2(\mathbf{Z})$. We have $\mathbf{Q}_f = \mathbf{Q}$ and $\lambda = l$ for some prime number l. For every $r \geq 1$, let ζ_{l^r} be a primitive l^r-th root of unity in $\overline{\mathbf{Q}}$. We can construct $\rho_{f,l}$ using the l-adic cyclotomic character

$$\chi_l\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gal}\bigl(\mathbf{Q}(\{\zeta_{l^r} \mid r \geq 1\})/\mathbf{Q}\bigr) \xrightarrow{\sim} \mathbf{Z}_l^\times \rightarrowtail \mathbf{Q}_l^\times,$$

which is defined uniquely by the equation

$$\sigma(\zeta_{l^r}) = \zeta_{l^r}^{\chi_l(\sigma) \bmod l^r} \quad \text{for all } r \geq 1 \text{ and all } \sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}).$$

It follows from the definition of χ_l that for every prime number $p \neq l$, we have

$$\chi_l(\mathrm{Frob}_p) = p \in \mathbf{Z}_l^\times.$$

(Note that $\chi_l(\mathrm{Frob}_l)$ is not defined.) This implies that up to conjugation in $\mathrm{GL}_2(\mathbf{Q}_l)$, the Galois representation attached to $f = E_k$ is

$$\rho_{f,l} = \begin{pmatrix} 1 & 0 \\ 0 & \chi_l^{k-1} \end{pmatrix}.$$

The construction that Deligne used in [7] to attach l-adic representations to cusp forms is much more difficult than for Eisenstein series. The main ingredients are:

- Hecke correspondences on $\mathrm{X}_1(n)$ (see § 2.3);
- the Eichler–Shimura isomorphism between modular forms and the cohomology of certain "local systems" on $\mathrm{X}_1(n)$, which can equivalently be interpreted as group cohomology (see Gabor Wiese's lectures);
- the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the l-adic étale cohomology of these local systems, or on the l-adic Tate module of the Jacobian of $\mathrm{X}_1(n)$ in the case of weight 2;
- the Eichler–Shimura congruence relation linking the Hecke operator T_p in characteristic p with the Frobenius map (see § 3.6 below for the statement in the context of the Jacobian of $\mathrm{X}_1(n)$).

3.3. Reduced representations

Using the compactness of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, one can show that after a suitable change of basis, the matrix $\rho_{f,\lambda}(\sigma)$ has algebraically integral coefficients for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. One can then reduce modulo λ and obtain a representation with values in $\mathrm{GL}_2(\mathbf{F}_\lambda)$, where $\mathbf{F}_\lambda$ is the residue field of λ. This representation in general depends on the choice of basis, but it becomes unique (up to conjugation) after an operation called *semi-simplification*. Finally, one can show that for the existence of this reduced representation, it is not strictly necessary that f be an eigenform; it suffices that f is an eigenform "modulo λ". The general statement about Galois representations attached to eigenforms over finite fields is the following.

Theorem 3.3 (Eichler, Shimura, Igusa, Deligne, Serre). *Let $f \in \mathrm{M}_k(n)(\mathbf{F})$ be a normalised Hecke eigenform of weight k for the group $\Gamma_1(n)$ with coefficients $a_m(f)$ in a finite field $\mathbf{F}$ of characteristic l and with character $\epsilon\colon (\mathbf{Z}/n\mathbf{Z}) \to \mathbf{F}^\times$. There exists a continuous group homomorphism*

$$\bar\rho_f\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{F})$$

with the following properties:

(1) $\bar\rho_f$ is unramified outside nl;

(2) for every prime number $p \nmid nl$, the characteristic polynomial of $\bar\rho_f(\sigma_p)$, where σ_p is a Frobenius element at p, equals $t^2 - a_p(f)t + \epsilon(p)p^{k-1}$.

Furthermore, this $\bar\rho_f$ is odd and unique up to conjugation in $\mathrm{GL}_2(\mathbf{F})$.

Let us explain this theorem in a slightly different way than we did for Theorem 3.1. There is a unique finite Galois extension $K_f/\mathbf{Q}$ such that the homomorphism $\bar\rho_f$ factors as

$$\bar\rho_f\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \twoheadrightarrow \mathrm{Gal}(K_f/\mathbf{Q}) \rightarrowtail \mathrm{GL}_2(\mathbf{F}).$$

The first condition means that every prime number $p \nmid nl$ is unramified in the field extension $K_f/\mathbf{Q}$. Let $\mathfrak{p}$ be a prime of K lying over p, let $k(\mathfrak{p})$ denote its residue field, and let $G_\mathfrak{p} \subseteq \mathrm{Gal}(K_f/\mathbf{Q})$ be the decomposition group at $\mathfrak{p}$. The fact that $\mathfrak{p}$ is unramified implies that $G_\mathfrak{p}$ maps isomorphically to $\mathrm{Gal}(k(\mathfrak{p})/\mathbf{F}_p)$. The element $\sigma_p \in G_\mathfrak{p} \subseteq \mathrm{Gal}(K_f/\mathbf{Q})$ is called a Frobenius element at p. It depends on the choice of $\mathfrak{p}$, but its conjugacy class in $\mathrm{Gal}(K_f/\mathbf{Q})$ is independent of this choice. The characteristic polynomial of $\bar\rho_f(\sigma_p)$ is invariant under conjugation of σ_p and hence only depends on p. Condition (2), which prescribes what this characteristic polynomial should be, therefore makes sense.

To simplify the presentation, we will assume in what follows that $\bar\rho_f$ is *absolutely irreducible*, i.e. that the representation $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}})$ obtained by base extension to an algebraic closure $\overline{\mathbf{F}}$ of $\mathbf{F}$ is irreducible.

Remark. The assumption that $\bar\rho_f$ is absolutely irreducible implies that f is a cusp form, because the Galois representations attached to Eisenstein series are direct sums of two characters. Moreover, using the fact that $\bar\rho_f$ is odd, one can show that if $l > 2$, then ρ_f is absolutely irreducible as soon as it is irreducible.

3.4. Jacobian varieties

Let X be a smooth projective curve over a field K. We recall some terminology on divisors and divisor classes.

A *divisor* on X is a finite formal sum of points on X, with integral (possibly negative) coefficients. The *degree* of a divisor is the sum of these coefficients. Any non-zero rational function on X has a divisor associated to it; such divisors are called *principal divisors*. Two divisors are *linearly equivalent* if their difference is a principal divisor. A *divisor class* is a linear equivalence class of divisors. For every extension L/K, there is a notion of *L-rational divisors* and of *L-rational divisor classes*.

There exists an *Abelian variety* (projective variety with a group structure) $\mathrm{Jac}\,X$ with the following property: for every field extension L/K, there is a bijection between the group $(\mathrm{Jac}\,X)(L)$ and the group of L-rational linear equivalence classes divisor of degree 0 on X. The dimension of $\mathrm{Jac}\,X$ is equal to the genus of X.

Let us now suppose that K is a subfield of $\mathbf{C}$. Let g be the genus of X. One has the following complex analytic description of $(\mathrm{Jac}\,X)(\mathbf{C})$ as the quotient of a complex vector space of dimension g by a lattice of rank $2g$:

$$(\mathrm{Jac}\,X)(\mathbf{C}) = \mathrm{H}^0(X(\mathbf{C}), \Omega^1_{X(\mathbf{C})})/\mathrm{H}_1(X(\mathbf{C}), \mathbf{Z}).$$

Let $f\colon Y \to X$ be a non-constant morphism of smooth projective curves. To f we can associate two maps between Jacobian varieties. First, we have the covariant *Albanese map*

$$h_*\colon \mathrm{Jac}\,Y \to \mathrm{Jac}\,X$$

sending the class of a divisor D on Y to the class of the push-forward of D by h. Second, we have the contravariant *Picard map*

$$h^*\colon \mathrm{Jac}\,X \to \mathrm{Jac}\,Y$$

sending the class of a divisor D on X to the class of the pull-back of D by h.

3.5. The Jacobian of $\mathrm{X}_1(n)$

Let $n \geq 4$ be an integer. We introduce the abbreviation

$$\mathrm{J}_1(n) = \mathrm{Jac}(\mathrm{X}_1(n)).$$

For every $m \geq 1$, we use the diagram (2.4) to define an endomorphism

$$T_m = (q_2)_* \circ (q_1)^* \colon \mathrm{J}_1(n) \to \mathrm{J}_1(n).$$

Similarly, for every $d \in (\mathbf{Z}/n\mathbf{Z})^\times$ we use (2.5) to define an automorphism

$$\langle d \rangle = (r_d)_* \colon \mathrm{J}_1(n) \xrightarrow{\sim} \mathrm{J}_1(n).$$

Definition. The *Hecke algebra* acting on $\mathrm{J}_1(n)$ is the commutative subring

$$\mathbf{T}(J_1(n)) \subseteq \mathrm{End}(\mathrm{J}_1(n))$$

generated as a $\mathbf{Z}$-algebra by all the T_m for $m \geq 1$ (or equivalently by the T_p for p prime and the $\langle d \rangle$ for $d \in (\mathbf{Z}/n\mathbf{Z})^\times$).

3.6. The Eichler–Shimura congruence relation

Let p be a prime number not dividing n. The fact that the modular curve $\mathrm{X}_1(n)$ has good reduction at p implies that its Jacobian $\mathrm{J}_1(n)$ has good reduction at p. Let Frob_p denote the *Frobenius endomorphism* of $\mathrm{J}_1(n)_{\mathbf{F}_p}$; this acts on a point by raising its coordinates to the p-th power. Let Ver_p denote the *Verschiebung**; this is the unique endomorphism of $\mathrm{J}_1(n)_{\mathbf{F}_p}$ satisfying

$$\mathrm{Frob}_p \circ \mathrm{Ver}_p = \mathrm{Ver}_p \circ \mathrm{Frob}_p = p. \tag{3.1}$$

In the ring $\mathrm{End}(\mathrm{J}_1(n)_{\mathbf{F}_p})$, we have the *Eichler–Shimura (congruence) relation*

$$T_p = \mathrm{Frob}_p + \langle p \rangle \, \mathrm{Ver}_p\,;$$

see Diamond and Im [10, § 8.5 and § 10.2]. Multiplying by Frob_p and using (3.1), we obtain

$$\mathrm{Frob}_p^2 - T_p \, \mathrm{Frob}_p + \langle p \rangle p = 0.$$

Moreover, if l is a prime number different from p, then the Tate module

$$\mathbf{Q}_l \otimes_{\mathbf{Z}_l} \varprojlim_r \mathrm{J}_1(n)(\bar{\mathbf{F}}_p)[l^r]$$

is a free module of rank 2 over $\mathbf{Q}_l \otimes \mathbf{T}(\mathrm{J}_1(n))$, and the characteristic polynomial of Frob_p on this space equals $t^2 - T_p t + \langle p \rangle p \in \mathbf{T}(\mathrm{J}_1(n))[t]$; see Diamond and Im [10, § 12.5].

3.7. Realisation of ρ_f *in the Jacobian of a modular curve*

Let f be a Hecke eigenform of weight k for the group $\Gamma_1(n)$ with coefficients in a finite field $\mathbf{F}$ of characteristic l.

Theorem 3.4 (see Serre [23, 2.7, remarque (2)] (without proof) or Edixhoven [12, Theorem 3.4]). *There exists an eigenform $\tilde{f}$ for the group $\Gamma_1(n)$ with coefficients in $\mathbf{F}$, but of weight $\tilde{k}$ with $1 \leq \tilde{k} \leq l+1$, such that*

$$\rho_f = \chi_l^i \otimes \rho_{\tilde{f}}$$

for some integer i.

On the level of modular forms, it means that

$$\begin{aligned} a_p(f) &= (p \bmod l)^i a_p(\tilde{f}) \quad (p \nmid nl \text{ prime}),\\ \epsilon_f(d) &= (p \bmod l)^{2i} \epsilon_{\tilde{f}}(d) \quad (d \in (\mathbf{Z}/n\mathbf{Z})^\times). \end{aligned}$$

* German for “shift” or “displacement”.

On the level of Galois representations, it means that if σ_p denotes a Frobenius element at a prime number $p \nmid nl$, we have

$$\rho_f(\sigma_p) = (p \bmod l)^i \rho_{\tilde{f}}(\sigma_p)$$

In case $\tilde{k} = 1$, we can find an eigenform of weight l for $\Gamma_1(n)$ over a quadratic extension $\mathbf{F}'$ of $\mathbf{F}$ whose associated Galois representation is isomorphic to $\mathbf{F}' \otimes_{\mathbf{F}} \rho_f$ [12, proof of Proposition 2.7].

After replacing f by $\tilde{f}$, we may therefore assume that the weight k of f satisfies

$$2 \le k \le l + 1.$$

It turns out that we now have to distinguish between two cases: the weight equals 2 or is greater than 2. Let us write

$$n' = \begin{cases} n & \text{if } k = 2; \\ nl & \text{if } 2 < k \le l + 1. \end{cases}$$

Theorem 3.5 (Gross [14, Proposition 11.8]; see also [13, Theorem 2.5.7] or [1, § I.3.6]). *There exists a ring homomorphism*

$$e_f \colon \mathbf{T}(\mathrm{J}_1(n')) \to \mathbf{F}$$

that maps T_m to $a_m(f)$ for all $m \ge 1$ and $\langle d \rangle$ to $\epsilon(d \bmod n) \cdot (d \bmod l)^{k-2}$; the last factor is understood to be 1 if $k = 2$.

We may assume that e_f is surjective. Let $\mathfrak{m}_f$ be the kernel of e_f. Then the set

$$V_f = \mathrm{J}_1(n')(\overline{\mathbf{Q}})[\mathfrak{m}_f].$$

has a natural structure of $\mathbf{F}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module. One can show that V_f is non-zero. Moreover, using the Eichler–Shimura relation, Čebotarev's density theorem, and a theorem due to Boston, Lenstra and Ribet, one can show that V_f is a direct sum of copies of the representation ρ_f. (Here we use the assumption that ρ_f is absolutely irreducible.)

Theorem 3.6 (Mazur, Ribet, Gross, Edixhoven, Buzzard, Wiese [29]).

(1) If ρ_f is ramified at l, or if ρ_f is unramified at l and a Frobenius element at l does not act as a scalar, then the $\mathbf{F}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module is isomorphic to ρ_f.

(2) If ρ_f is unramified at l, a Frobenius element at l does act as a scalar, and ρ_f arises from a form of weight 1 (the last condition follows from the others if $l > 2$), then $\mathrm{J}_1(n')[\mathfrak{m}_f](\overline{\mathbf{Q}})$ is a direct sum of at least two copies of ρ_f.

Roughly speaking, the conclusion of this section is that up to twisting, the representation ρ_f occurs in the l-torsion of the Jacobian of a modular curve.

4. Computation of modular Galois representations

References: the book [13] edited by Edixhoven and Couveignes, and the author's thesis [1].

4.1. The question and the main result

Let f be a Hecke eigenform over a finite field $\mathbf{F}$. We are interested in the problem of "computing" the modular Galois representation

$$\rho_f \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{F})$$

attached to f.

Since $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is infinite, we first have to show that we can describe ρ_f using a finite amount of data. We recall that ρ_f factors as

$$\rho_f \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \twoheadrightarrow \mathrm{Gal}(K/\mathbf{Q}) \rightarrowtail \mathrm{GL}_2(\mathbf{F}),$$

where K is a finite Galois extension of $\mathbf{Q}$. This means that it suffices to compute the field K (which we can describe for example as the splitting field of an irreducible polynomial over $\mathbf{Q}$) together with an embedding of $\mathrm{Gal}(K/\mathbf{Q})$ into $\mathrm{GL}_2(\mathbf{F})$.

The above question has the following answer.

Theorem 4.1 (Edixhoven, Couveignes et al. [13] for $n = 1$; B. [1] for $n > 1$). *There exists an algorithm that, given*

- *positive integers n and k, with n square-free*
- *a finite field $\mathbf{F}$ of characteristic greater than k, and*
- *an eigenform f of weight k for $\Gamma_1(n)$ with coefficients in $\mathbf{F}$, given by the $a_m(f)$ for $0 \le m \le \frac{k}{12}(\mathrm{SL}_2(\mathbf{Z}) : \{\pm 1\}\Gamma_1(n))$,*

computes ρ_f in the form of the following data:

- *the finite Galois extension K of $\mathbf{Q}$ such that ρ_f factors as*

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \twoheadrightarrow \mathrm{Gal}(K/\mathbf{Q}) \rightarrowtail \mathrm{GL}_2(\mathbf{F}),$$

  given by the multiplication table of some $\mathbf{Q}$-basis $(b_1, \ldots, b_r)$ of K;
- *for every $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$, the matrix of σ with respect to the basis $(b_1, \ldots, b_r)$ and the element $\rho_f(\sigma) \in \mathrm{GL}_2(\mathbf{F})$,*

and that runs in time polynomial in k, n and $\#\mathbf{F}$.

Moreover, once ρ_f has been computed, one can compute $\rho_f(\mathrm{Frob}_p)$ using a deterministic algorithm in time polynomial in k, n, $\#\mathbf{F}$ and $\log p$.

Remark. The algorithm for $n = 1$ given in [13] is deterministic. For $n > 1$, only a probabilistic algorithm is known. The word 'probabilistic' here means that the output is guaranteed to be correct, but that the running time depends on random choices made during the execution of the algorithm. It is the *expected* running time (taken with respect to these random choices, *not* with respect to the possible inputs) that is polynomial in the length of the input.

Remark. The assumption that n is square-free is made for technical reasons and will probably be removed in the near future.

4.2. Reduction to computation of vector space schemes

Using the techniques described in the previous talk, we reduce the problem to computing Galois representations of the form

$$\rho_{J[\mathfrak{m}]}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}_{\mathbf{F}}\, J[\mathfrak{m}](\overline{\mathbf{Q}})$$

where J is the Jacobian of the modular curve $\mathrm{X}_1(n)$ for some n, $\mathfrak{m}$ is a maximal ideal of the Hecke algebra acting on J and $\mathbf{F}$ is the residue field of $\mathfrak{m}$.

Computing the representation ρ_f comes down to computing $J[\mathfrak{m}]$ as a two-dimensional $\mathbf{F}$-vector space scheme over $\mathbf{Q}$. We do this in the following way. We choose an embedding of finite $\mathbf{Q}$-schemes

$$\iota\colon J[\mathfrak{m}] \to \mathbf{A}^1_{\mathbf{Q}}.$$

This gives a description of $J[\mathfrak{m}]$ as $\mathrm{Spec}\,\mathbf{Q}[x]/P$, for some $P \in \mathbf{Q}[x]$. The $\mathbf{F}$-vector space structure on $\mathrm{Spec}\,\mathbf{Q}[x]/P$ is given (in the style of Hopf algebras) by ring homomorphisms

$$\alpha\colon \mathbf{Q}[x]/(P) \to \mathbf{Q}[x_1, x_2]/(P(x_1), P(x_2))$$

and

$$\mu_c\colon \mathbf{Q}[x]/(P) \to \mathbf{Q}[x]/(P) \quad \text{for all } c \in \mathbf{F}$$

describing addition and scalar multiplication, respectively.

The strategy for computing $\rho_{J[\mathfrak{m}]}$ consists of two parts:

(1) approximate P, α and the μ_c by computing them modulo p for sufficiently many small prime numbers p;

(2) using height bounds, reconstruct the data over $\mathbf{Q}$ from these approximations.

Remark. Instead of computing the data modulo p for many small prime numbers p, one can try to take approaches based on complex or p-adic approximation of the data. We will not describe these variants.

4.3. Computing in Jacobians of curves of modular curves

Let us now briefly describe one of the essential tools used in the algorithm, namely a collection of algorithms for computing with divisors on curves over finite fields. The general set-up, which gives the asymptotically fastest known algorithms for computing with divisors on general curves, was developed by Khuri-Makdisi in [16] and [17]. Various extensions were developed by the author [3] to deal with finite morphisms between curves and with operations specific to curves over finite fields.

For concreteness, we will restrict to modular curves over finite fields. Let $n \ge 11$ be an integer, and let $\mathrm{X}_1(n)$ be the compactified modular curve over $\mathbf{Z}[1/n]$ classifying elliptic curves with a point of order n. (The reason for the restriction $n \ge 11$ is both that there are small technical difficulties for $n \le 4$, and that $\mathrm{X}_1(n)$ has genus 0 for all $n \le 10$ and is therefore less interesting.) Let p a prime number not dividing n. We want to describe (very roughly) how one can compute in the Jacobian $\mathrm{J}_1(n)_{\mathbf{F}_p}$ of the curve $\mathrm{X}_1(n)_{\mathbf{F}_p}$.

For $i \ge 0$, let V_i denote the vector space of modular forms of weight $2i$ for $\Gamma_1(n)$ with coefficients in $\mathbf{F}_p$. The V_i are finite-dimensional, and there are multiplication maps

$$V_i \times V_j \to V_{i+j}.$$

The geometric interpretation is that V_i consists of the global sections of ω^{2i}, where ω is the line bundle of modular forms on $\mathrm{X}_1(n)$. We will only need finitely many V_i; in fact, for our purposes it will be enough to know the V_i for $1 \le i \le 7$ together with the multiplication maps between them.

Let D be an effective divisor on $\mathrm{X}_1(n)$. Then for i sufficiently large, D is uniquely determined by the subspace of V_i consisting of forms (of weight $2i$) that vanish on D. In this way, we get a way of representing effective divisors. We represent divisors of degree 0 as differences of two effective divisors of the same degree. Using (bi)linear algebra on subspaces of the V_i, one can perform various operations such as testing for linear equivalence, and addition and subtraction of divisors and divisor classes.

For us, the most relevant result is that one can compute the l-torsion of $\mathrm{J}_1(n)$ over a finite extension of $\mathbf{F}_p$.

Theorem 4.2. *There exists a probabilistic algorithm that, given a positive integer n, a finite field k of characteristic $p \nmid n$ and cardinality q, a prime number $l \ne p$, computes an $\mathbf{F}_l$-basis for $\mathrm{J}_1(n)(k)[l]$ in expected time polynomial in n, p, $[k : \mathbf{F}_p]$ and l.*

4.4. Computing coefficients of modular forms

The research project leading to [13] and [1] was originally motivated by a question that René Schoof asked to Bas Edixhoven in 1995: can one evaluate Ramanujan's τ-function at a prime number p in time polynomial in $\log p$? Ramanujan's τ-function is defined by the following equality of power series:

$$\sum_{m=1}^{\infty} \tau(m) q^m = q \prod_{n=1}^{\infty} (1 - q^n)^{24}.$$

This is the q-expansion of the modular form Δ, the unique normalised cusp form of weight 12 for the full modular group $\mathrm{SL}_2(\mathbf{Z})$.

More generally, one can ask the question how fast one can compute the m-th coefficient of f for a given positive integer m and a given modular form f with coefficients in some number field. One can use Theorem 4.1 together with some analytic number theory to accomplish this, assuming the generalised Riemann hypothesis. The result is as follows.

Theorem 4.3 (Edixhoven and Couveignes [13] for $n = 1$; B. [2, Theorem 1.1] for $n > 1$). *There exists an algorithm that, given*

- *positive integers n and k, with n square-free,*
- *a number field K,*
- *a modular form f of weight k for $\Gamma_1(n)$ over K, and*
- *a positive integer m in factored form,*

computes $a_m(f)$, and whose running time is bounded by a polynomial in the length of the input under the Riemann hypothesis for zeta functions of number fields.

Remark. As in Theorem 4.1, a deterministic algorithm is known for $n = 1$, but only a probabilistic algorithm for $n > 1$.

Remark. One can hardly expect to remove the assumption that m is given in factored form. Namely, the multiplicative relations between the coefficients of an eigenform imply that if the theorem were true without this assumption, then we would be able to factor integers (at least products of two prime numbers) in polynomial time.

4.5. Example: sums of squares

It is a classical problem to determine in how many ways a given positive integer m can be written as a sum of k squares for given k. In other words, the problem is to evaluate the function

$$r_k(m) = \#\{(x_1, \ldots, x_k) \in \mathbf{Z}^k \mid x_1^2 + \cdots + x_k^2 = m\}.$$

This function is famously related to modular forms. We consider Jacobi's theta series

$$\theta = \sum_{m \in \mathbf{Z}} q^{m^2} = 1 + 2\sum_{m=1}^{\infty} q^{m^2} \in \mathbf{Z}[[q]].$$

An easy combinatorial argument shows that for every positive integer k, we have

$$\theta^k = \sum_{m=0}^{\infty} r_k(m) q^m.$$

It is known that θ is (the q-expansion of) a modular form of weight 1/2 for the group $\Gamma_1(4)$; this is essentially Poisson's summation formula. This implies that θ^k is a modular form of weight $k/2$ for $\Gamma_1(4)$. The analogue of Theorem 4.3 with $n = 4$ is also true, and we obtain the following result.

Theorem 4.4 (B. [2, Theorem 6.5]). *There exists a probabilistic algorithm that, given an even positive integer k and a positive integer m in factored form, computes $r_k(m)$, and that runs in time polynomial in k and $\log m$ under the Riemann hypothesis for zeta functions of number fields.*

References

[1] P. J. Bruin, *Modular curves, Arakelov theory, algorithmic applications.* Ph. D. thesis, Universiteit Leiden, 2010. Available on the web: `http://hdl.handle.net/1887/15915`.

[2] P. J. Bruin, Computing coefficients of modular forms. *Publications mathématiques de Besançon*, année 2011 (actes de la conférence *Théorie des nombres et applications*, CIRM, Marseille, 30 novembre–4 décembre 2009), 19–36.

[3] P. J. Bruin, Computing in Picard groups of projective curves over finite fields. *Mathematics of Computation*, to appear.

[4] J. W. S. Cassels, *Lectures on Elliptic Curves.* London Mathematical Society Student Texts **24**. Cambridge University Press, 1991.

[5] B. Conrad, Arithmetic moduli of generalized elliptic curves. *Journal de l'Institut de Mathématiques de Jussieu* **6** (2007), no. 2, 209–278.

[6] B. Conrad, *Modular Forms and the Ramanujan Conjecture.* Cambridge University Press, to appear.

[7] P. Deligne, Formes modulaires et représentations l-adiques. Séminaire Bourbaki, 21e année (1968/1969), exposé 355. Lecture Notes in Mathematics **179**, 139–172. Springer-Verlag, Berlin/Heidelberg/New York, 1971.

[8] P. Deligne et M. Rapoport, Les schémas de modules de courbes elliptiques. In: P. Deligne and W. Kuyk (editors), *Modular Functions of One Variable II* (Proceedings of the International Summer School, University of Antwerp, 1972). Lecture Notes in Mathematics **349**, 143–316. Springer-Verlag, Berlin/Heidelberg, 1973.

[9] P. Deligne et J.-P. Serre, Formes modulaires de poids 1. *Annales scientifiques de l'É.N.S. (4^e^ série)* **7** (1974), no. 4, 507–530.

[10] F. Diamond and J. Im, Modular forms and modular curves. In: V. Kumar Murty (editor), *Seminar on Fermat's Last Theorem* (Fields Institute for Research in Mathematical Sciences, Toronto, ON, 1993–1994), 39–133. CMS Conference Proceedings **17**. American Mathematical Society, Providence, RI, 1995.

[11] F. Diamond and J. Shurman, *A First Course in Modular Forms*. Springer-Verlag, Berlin/Heidelberg/New York, 2005.

[12] S. J. Edixhoven, The weight in Serre's conjectures on modular forms. *Inventiones mathematicae* **109** (1992), 563–594.

[13] S. J. Edixhoven and J.-M. Couveignes (with J. G. Bosman, R. S. de Jong and F. Merkl), *Computational aspects of modular forms and Galois representations*. Annals of Mathematics Studies **176**. Princeton University Press, Princeton/Oxford, 2011.

[14] B. H. Gross, A tameness criterion for Galois representations associated to modular forms (mod p). *Duke Mathematical Journal* **61** (1990), no. 2, 445–517.

[15] N. M. Katz and B. Mazur, *Arithmetic Moduli of Elliptic Curves*. Annals of Mathematics Studies **108**. Princeton University Press, Princeton, NJ, 1985.

[16] K. Khuri-Makdisi, Linear algebra algorithms for divisors on an algebraic curve. *Mathematics of Computation* **73** (2004), no. 245, 333–357.
Available on the web: `http://arxiv.org/abs/math.NT/0105182`.

[17] K. Khuri-Makdisi, Asymptotically fast group operations on Jacobians of general curves. *Mathematics of Computation* **76** (2007), no. 260, 2213–2239.
Available on the web: `http://arxiv.org/abs/math.NT/0409209`.

[18] B. Mazur, Modular curves an the Eisenstein ideal. *Publications mathématiques de l'I.H.É.S.* **47** (1977), 33–186.

[19] J. S. Milne, *Elliptic Curves*. BookSurge Publishers, 2006.
Available on the web: `http://www.jmilne.org/math/Books/ectext.html`.

[20] T. Miyake, *Modular Forms*. Springer Monographs in Mathematics. Springer-Verlag, Berlin/Heidelberg, 2006.

[21] K. A. Ribet, On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms. *Inventiones mathematicae* **100** (1990), 431–476.

[22] K. A. Ribet and W. A. Stein, *Lectures on modular forms and Hecke operators*. Available on the web: `http://wstein.org/books/ribet-stein/main.pdf`.

[23] J-P. Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Mathematical Journal* **54** (1987), 179–230. (= Œuvres, IV, **143**. Springer-Verlag, Berlin/Heidelberg, 2000.)

[24] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Forms*. Iwanami Shoten, Tokyo, and Princeton University Press, 1971; Princeton University Press, 1994.

[25] J. H. Silverman, *The Arithmetic of Elliptic Curves*. Graduate Texts in Mathematics **106**. Springer-Verlag, Berlin/Heidelberg/New York, 1992.

[26] J. H. Silverman, *Advanced Topics in the Theory of Elliptic Curves*. Graduate Texts in Mathematics **151**. Springer-Verlag, Berlin/Heidelberg/New York, 1994.

[27] J. H. Silverman and J. T. Tate, *Rational Points on Elliptic Curves*. Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1992.

[28] J. T. Tate, The arithmetic of elliptic curves. *Inventiones mathematicae* **23** (1974), 179–206.

[29] G. Wiese, Multiplicities of Galois representations of weight one. With an appendix by N. Naumann. *Algebra and Number Theory* **1** (2007), no. 1, 67–85.