Modularity for elliptic curves and beyond

Jack A. Thorne

September 1, 2017

Contents

1	Introduction	1
2	Lectures	1
3	Exercises	17

1 Introduction

The aim of these notes is to give an introduction to the notion of modularity of elliptic curves and related objects. This is a vast topic, and we can barely scratch the surface here. We therefore focus on the basic definitions, and their consequences for arithmetic and for the properties of L-functions. We pass over in silence the question of how modularity theorems are actually proved; and we can mention only very briefly the theory of automorphic representations (and the foundational results in the representation theory of real and p-adic groups) that gives the deepest understanding of the picture we sketch here. For the reader who wishes to go further, we have included references to the wider literature at the end of each lecture.

2 Lectures

2.1 Lecture 1

Let E be an elliptic curve over $\mathbb Q$ of conductor $N=N_E.$ Its L-function

$$L(E,s) \doteq \prod_{p \text{ prime}} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

is a function of a complex variable s, defined by an infinite product which converges absolutely in the region $\text{Re } s > 3/2.^1$ It admits an analytic continuation to the whole complex plane, and satisfies the functional

¹Here we use \doteq to denote equality up to finite many factors in the Euler product, namely those corresponding to primes p where E has bad reduction.

equation $\Lambda(E,s) = \pm \Lambda(E,2-s)$, where by definition

$$\Lambda(E,s) = N^{s/2} (2\pi)^{-s} \Gamma(s) L(E,s).$$

The goal of this lecture is to explain what this picture has to do with modular forms.

Let $\mathbb{H} = \{ \tau \in \mathbb{C} \mid \text{im } \tau > 0 \}$ denote the usual complex upper half plane. The group $GL_2(\mathbb{R})^+$ (real matrices with positive determinant) acts transitively on \mathbb{H} by Möbius transformations:

$$\left(\begin{array}{cc} a & b \\ c & d \end{array}\right)\tau = \frac{a\tau + b}{c\tau + b}.$$

Let $\Gamma = \operatorname{SL}_2(\mathbb{Z}) \subset \operatorname{GL}_2(\mathbb{R})^+$; then the group Γ acts properly and discontinuously on \mathbb{H} . (By definition, this means that for any $\tau_1, \tau_2 \in \mathbb{H}$, there exist open neighbourhoods U_1 of τ_1 and U_2 of τ_2 in \mathbb{H} with the following property: for any $\gamma \in \Gamma$, we have $\gamma(U_1) \cap U_2 \neq \emptyset \Rightarrow \gamma(\tau_1) = \tau_2$.)

For any $N \geq 1$, we define the congruence subgroup

$$\Gamma_0(N) = \left\{ \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in \Gamma \mid c \equiv 0(N) \right\}.$$

The quotient $Y_0(N) = \Gamma_0(N) \backslash \mathbb{H}$ is a Hausdorff topological space, and in fact has a natural structure of Riemann surface, which we describe in the exercises. This Riemann surface can be compactified by adding finitely many 'cusps' as follows: let $\mathbb{H}_{\infty} = \mathbb{H} \sqcup \mathbb{P}^1(\mathbb{Q})$. The group Γ acts on \mathbb{H}_{∞} in a natural way extending its action on \mathbb{H} . We give \mathbb{H}_{∞} the topology where \mathbb{H} is an open subspace and, for each element $\gamma \in \Gamma$, the point $\gamma(\infty)$ has a basis of open neighbourhoods of the form $\gamma(U_y \cup \{\infty\}) = \gamma(\{\tau \in \mathbb{H} \mid \operatorname{Im} \tau > y\} \cup \infty)$. This describes the topology, since Γ acts transitively on $\mathbb{P}^1(\mathbb{Q})$!

One can show that $X_0(N) = \Gamma_0(N) \backslash \mathbb{H}_{\infty}$ is a compact Hausdorff space, and has a natural structure of connected compact Riemann surface. We write $S_2(\Gamma_0(N), \mathbb{C})$ for the vector space $H^0(X_0(N), \Omega^1_{X_0(N)})$; it is canonically identified with the usual space of cuspidal holomorphic modular forms of weight 2 and level $\Gamma_0(N)$. More precisely, if $\omega \in H^0(X_0(N), \Omega^1_{X_0(N)})$, then the pullback of ω to \mathbb{H} can be written as $F(\tau)d\tau$, where $F: \mathbb{H} \to \mathbb{C}$ is a holomorphic function. Those who are familiar with the definitions can check that $F(\tau)$ is cuspidal holomorphic modular form of weight 2 and level $\Gamma_0(N)$, and conversely that any such function $F(\tau)$ determines a $\Gamma_0(N)$ -invariant holomorphic differential on \mathbb{H} which descends to an element of $H^0(X_0(N), \Omega^1_{X_0(N)})$.

It is a fact that $X_0(N)$ can be defined canonically as an algebraic curve over \mathbb{Q} . We now change notation and write $X_0(N)$ for this algebraic curve over \mathbb{Q} (and $Y_0(N) \subset X_0(N)$ for the open subvariety, also defined over \mathbb{Q} , which is the complement of the cusps). The existence of this model for $X_0(N)$ is a consequence of its interpretation as a moduli space for elliptic curves. The starting point for this is the following lemma.

Lemma 2.1. The map $\tau \in \mathbb{H} \mapsto (E_{\tau}, C_{\tau}) = (\mathbb{C}/(\mathbb{Z} \oplus \mathbb{Z}\tau), (\frac{1}{N}\mathbb{Z} \oplus \mathbb{Z}\tau)/(\mathbb{Z} \oplus \mathbb{Z}\tau))$ determines a bijection between the following two sets:

- The set $\Gamma_0(N)\backslash \mathbb{H}$.
- The set of equivalence classes of pairs (E,C), where E is an elliptic curve over \mathbb{C} and $C \subset E$ is a cyclic subgroup of order N. Two such pairs are said to be equivalent if there exists an isomorphism $f: E \to E'$ of elliptic curves over \mathbb{C} such that f(C) = C'.

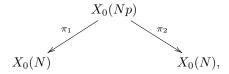
The curve $Y_0(N)$ over \mathbb{Q} is a coarse moduli space for pairs (E, C), where E is an elliptic curve and $C \subset E$ is a cyclic subgroup of order N. For example, it has the property that for any field extension K/\mathbb{Q} , the set $Y_0(N)(K)$ is in bijection with the set of equivalence classes of pairs (E, C), where E is an elliptic curve

over K and $C \subset E$ is a cyclic subgroup of order N. Two such pairs (E,C) and (E',C') are said to be equivalent if there exists an isomorphism $f: E_{\overline{K}} \to E'_{\overline{K}}$ (defined over an algebraic closure \overline{K} of K) such that $f(C_{\overline{K}}) = C'_{\overline{K}}$. The curve $X_0(N)$ can be interpreted as a (coarse) moduli space of 'generalized elliptic curves': the cusps correspond to degenerations of elliptic curves to so-called Néron polygons, which have a toric connected component.

The Jacobian $J_0(N) = \operatorname{Pic}^0 X_0(N)$ of $X_0(N)$ is an abelian variety over \mathbb{Q} of dimension equal to the genus of $X_0(N)$. The introduction of $J_0(N)$ allows us to define what it means for an elliptic curve to be modular.

Definition 2.2. Let E be an elliptic curve over \mathbb{Q} of conductor $N = N_E$. We say that E is modular if there exists a surjective homomorphism $\pi: J_0(N) \to E$.

We want to explain the consequences of this definition for the L-function L(E,s). The key is a set of operators, called the Hecke operators, which act both as endomorphisms of the vector space $S_2(\Gamma_0(N), \mathbb{C})$ and as endomorphisms of the Jacobian $J_0(N)$. For every prime p not dividing N, we can define an endomorphism T_p of $J_0(N)$, called the pth Hecke operator. It can be defined using the functorial properties of the Jacobian as follows. There is a diagram of compact Riemann surfaces:



where these maps are given on \mathbb{H} by the formulae $\pi_1(\tau) = p\tau$ and $\pi_2(\tau) = \tau$, respectively. We set $T_p = \pi_{2,*} \circ \pi_1^* \in \operatorname{End}(J_0(N))$. These maps can be described also in terms of the moduli interpretation of $Y_0(N)$ as follows: let us think of $Y_0(Np)$ as parameterizing tuples (E, C_N, C_p) , where $C_N \subset E$ is a cyclic subgroup of order N and C_p is a cyclic subgroup of order p, so $C_N \times C_p$ is a cyclic subgroup of order Np. Then $\pi_1(E, C_N, C_p) = (E/C_p, C_N + C_p/C_p)$, and $\pi_2(E, C_N, C_p) = (E, C_N)$.

The Hecke operators T_p allow us to make the link with L-functions. Let E be an elliptic curve over \mathbb{Q} , and suppose that there is a surjective homomorphism $\pi: J_0(N) \to E$.

Lemma 2.3 (Eichler–Shimura relation). Let p be a prime not dividing N. Then $\pi \circ T_p = [a_p] \circ \pi$. (Here $[n] \in \operatorname{End}_{\mathbb{Q}}(E)$ is the endomorphism 'multiplication by n'.)

Proof. In the exercises, we discuss how this follows from understanding the action of T_p on the mod p fibre of $J_0(N)$ (an abelian variety over \mathbb{Q} which has good reduction at the prime p). This uses the description of T_p in terms of its action on moduli.

Let $\omega_E \in H^0(E, \Omega_E^1)$ be a non-zero differential. The lemma implies that $\omega = i^* \pi^* \omega_E \in H^0(X_0(N), \Omega^1_{X_0(N)})$ is a simultaneous eigenvector for all of the operators T_p , with eigenvalue $a_p \in \mathbb{Z}$. The differential form ω can be represented as a holomorphic differential $F(\tau)d\tau$ on \mathbb{H} , which is invariant under the action of $\Gamma_0(N)$. In particular, it is invariant under the transformation $\tau \mapsto \tau + 1$, which corresponds to the action of the matrix

$$\left(\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right) \in \Gamma_0(N).$$

We find that $F(\tau)d\tau$ descends to a differential $f(q)\frac{dq}{q}$ on the unit disc $\{q\in\mathbb{C}\mid |q|<1\}$, where $q=e^{2\pi i\tau}$. This differential can be represented by its Taylor expansion $f(q)\frac{dq}{q}=\sum_{n\geq 1}b_nq^n\frac{dq}{q}$.

The following is a consequence of the explicit theory of Hecke operators on $S_2(\Gamma_0(N), \mathbb{C})$:

Theorem 2.4. We have $b_1 \in \mathbb{Q}^{\times}$. After rescaling ω_E so that $b_1 = 1$, we have $a_n = b_n$ for all $n \geq 1$, where $L(E, s) = \sum_{n \geq 1} a_n n^{-s}$.

Proof. See [DS05, Ch. 5]. The easiest case is the equality $a_p = b_p$ when $p \nmid N$: the action of T_p on $f(q) \frac{dq}{q}$ is given by the formula

$$T_p\left(\sum_{n\geq 1}b_nq^n\frac{dq}{q}\right) = \left(\sum_{n\geq 1}(b_{np} + pb_{n/p})q^n\right)\frac{dq}{q}.$$

Thus the equality $T_p f(q) \frac{dq}{q} = a_p f(q) \frac{dq}{q}$ implies, after looking at the first coefficient, that $a_p = b_p$.

Corollary 2.5. If E is modular, then L(E,s) has an analytic continuation to the whole complex plane and satisfies the expected functional equation.

Proof. Let $i^*\pi^*\omega_E = F(\tau)d\tau = f(q)\frac{dq}{q}$. The Mellin transform of $F(\tau)$ is defined to be the integral $\int_{t=0}^{\infty} F(it)t^s\frac{dt}{t}$. We have

$$\int_{t=0}^{\infty} F(it)t^{s} \frac{dt}{t} = \int_{t=0}^{\infty} \sum_{n>1} b_{n} e^{-2\pi nt} t^{s} \frac{dt}{t},$$

and this double integral/sum is absolutely convergent when $\operatorname{Re} s$ is sufficiently large. We can therefore reverse the order of integration to get

$$\sum_{n\geq 1} b_n \int_{t=0}^{\infty} e^{-2\pi nt} t^s \frac{dt}{t} = \sum_{n\geq 1} b_n n^{-s} (2\pi)^{-s} \Gamma(s).$$

Since $\Lambda(E,s) = (2\pi)^{-s}\Gamma(s)N^{s/2}L(E,s)$, we can rewrite this as

$$\Lambda(E,s) = \int_{t=0}^{\infty} F(it/\sqrt{N}) t^{s} \frac{dt}{t},$$

an identity valid whenever Re s is sufficiently large. We will prove the analytic continuation and functional equation of $\Lambda(E,s)$ at the same time. Consider the matrix

$$w_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Q}).$$

It normalises $\Gamma_0(N)$ and induces an involution $w_N: X_0(N) \to X_0(N)$, $\tau \mapsto -1/N\tau$. It is a fact that we have $w_N^*F(\tau)d\tau = -w_EF(\tau)d\tau$, where $w_E \in \{\pm 1\}$ is a sign which is called the *root number* of the elliptic curve E.

This implies that we have $F(\tau) = w_E F(-1/N\tau)/N\tau^2$. A simple manipulation gives

$$\Lambda(E,s) = \int_{t=1}^{\infty} F(it/\sqrt{N}) t^{s} \frac{dt}{t} + \int_{t=0}^{1} F(it/\sqrt{N}) t^{s} \frac{dt}{t} = \int_{t=1}^{\infty} F(it/\sqrt{N}) t^{s} \frac{dt}{t} + w_{E} \int_{t=1}^{\infty} F(it/\sqrt{N}) t^{2-s} \frac{dt}{t}.$$

Since $F(it/\sqrt{N})$ decays very rapidly as $t \to \infty$, these integrals converge for any value of s. The functional equation $\Lambda(E, 2 - s) = w_E \Lambda(E, s)$ is obvious from this expression.

Of course, the point of all this is the following theorem:

Theorem 2.6 (Wiles, Taylor–Wiles, Breuil–Conrad–Diamond–Taylor). Every elliptic curve E over \mathbb{Q} is modular.

Corollary 2.7. For any elliptic curve E over \mathbb{Q} , L(E,s) admits an analytic continuation to the complex plane and satisfies a functional equation there.

References

An excellent introductory source, which contains all the material discussed in this lecture, is Diamond and Shurman's textbook [DS05]. In particular, it includes a proof of the fundamental Eichler-Shimura relation. The fundamental reference describing the model $X_0(N)$ over \mathbb{Q} (and over \mathbb{Z}_p , when $p \nmid N$) is Deligne-Rapoport [DR73]; it is based on sophisticated algebro-geometric techniques.

The modularity Theorem 2.6 began life as the Taniyama–Shimura–Weil conjecture. The first serious progress was made by Wiles and Taylor [Wil95, TW95], who proved the modularity of all semistable elliptic curves over \mathbb{Q} by proving the first modularity lifting theorems. The proof of the theorem for all elliptic curves over \mathbb{Q} was completed by Breuil, Conrad, Diamond, and Taylor [BCDT01]. An excellent introduction to the techniques involved in the proof is the article of Darmon, Diamond, and Taylor [DDT94].

2.2 Lecture 2

Let K be a number field. We associate to K the following notation:

- \mathcal{O}_K is the ring of integers of K.
- \mathcal{M}_K is the set of places of K (i.e. equivalence classes of non-trivial absolute values).
- If $v \in \mathcal{M}_K$, then K_v is the completion of K at v.
- If $v \in \mathcal{M}_K$ is a finite (i.e. non-archimedean) place, corresponding to a prime ideal $\mathfrak{p}_v \subset \mathcal{O}_K$, then $\mathcal{O}_{K_v} \subset K_v$ is the ring of integers of K_v , $k(v) = \mathcal{O}_{K_v}/(\mathfrak{p}_v)$ is the residue field of \mathcal{O}_{K_v} , and $q_v = \#k(v)$ is the cardinality of k(v).
- $\mathbb{A}_K = \prod_v' K_v$ is the adele ring of K, and $\mathbb{A}_K^{\infty} = \prod_{v \nmid \infty}' K_v$ is its finite part. By definition, we have (as a ring)

$$\mathbb{A}_K^{\infty} = \{(x_v)_v \in \prod_{v \nmid \infty} K_v \mid \text{ for all but finitely many } v, x_v \in \mathcal{O}_{K_v}\}.$$

 \mathbb{A}_K^{∞} contains the ring $\widehat{\mathcal{O}}_K = \prod_{v \nmid \infty} \mathcal{O}_{K_v}$ (profinite completion of the ring of integers of K) as an open subring.

We would like to define what it means for an elliptic curve E over K to be modular, in a way extending the definition given last time in the case $K = \mathbb{Q}$. We defined an elliptic curve E over \mathbb{Q} to be modular if there was a surjective homomorphism $\pi: J_0(N) \to E$. It's not clear how to generalize this statement to other number fields since in general there is no analogue of $J_0(N)$!

We first explain what the analogue of a modular curve is over a general number field K. Fix an isomorphism $K \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. We let $X = \operatorname{GL}_2(K \otimes_{\mathbb{Q}} \mathbb{R})/(\operatorname{O}(2)^{r_1} \times \operatorname{U}(2)^{r_2})\mathbb{R}^{\times}$. Then $\operatorname{GL}_2(K \otimes_{\mathbb{Q}} \mathbb{R})$ acts transitively on X. If $K \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{R}$, then we have $X = \mathbb{H}$, the complex upper half plane; if $K \times_{\mathbb{Q}} \mathbb{R} = \mathbb{C}$, then $K = \mathbb{H}_3$ is the hyperbolic upper half space (see the exercises).

For any open compact subgroup $U \subset \mathrm{GL}_2(\mathbb{A}_K^{\infty})$, we set

$$Y_U = \operatorname{GL}_2(K) \backslash \operatorname{GL}_2(\mathbb{A}_K^{\infty}) \times X/U \cong \sqcup_{g \in \operatorname{GL}_2(K) \backslash \operatorname{GL}_2(\mathbb{A}_\infty^{\infty})/U} \Gamma_{g,U} \backslash X.$$

It is a fact that the set $GL_2(K)\backslash GL_2(\mathbb{A}_K^{\infty})/U$ is finite, and if $g \in GL_2(\mathbb{A}_K^{\infty})$, then $\Gamma_{g,U} = GL_2(K) \cap gUg^{-1}$ is a congruence subgroup of $GL_2(K)$, so that Y_U is a disjoint union of finitely many copies of X quotiented by

a congruence subgroup of $GL_2(K)$. In special cases, Y_U can be simpler. For example, if K has class number 1 and $U = GL_2(\widehat{\mathcal{O}}_K)$, then $Y_U = GL_2(\mathcal{O}_K) \setminus X$.

If $\mathfrak{n} \subset \mathcal{O}_K$ is a non-zero ideal, then we can define a congruence subgroup

$$U_0(\mathfrak{n}) = \left\{ \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in \operatorname{GL}_2(\widehat{\mathcal{O}}_K) \mid c \equiv 0 \bmod \mathfrak{n} \right\}.$$

We write $Y_0(\mathfrak{n}) = Y_{U_0(\mathfrak{n})}$. If $K = \mathbb{Q}$ and $\mathfrak{n} = (N)$, then $Y_0(N)$ is just the open modular curve considered in the previous lecture. More generally, if K is a totally real field, then $Y_0(\mathfrak{n})$ is a circle bundle over a Hilbert modular variety. However, when K has a complex place, then $Y_0(\mathfrak{n})$ admits no natural complex structure, so a fortiori no structure of an algebraic variety. In particular, it doesn't make sense to form its Jacobian (or Albanese variety, or Picard variety, etc.). This is what we mean when we say there is no analogue of $J_0(N)$.

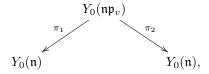
We therefore need to formulate the notion of modularity in a different way. We saw last time that given a modular parameterization $J_0(N) \to E$, the pullback $i^*\pi^*\omega_E \in H^0(X_0(N), \Omega^1_{X_0(N)})$ is an eigenvector for all the Hecke operators T_p . The Hecke operators also act on the singular cohomology of $X_0(N)$, and the isomorphism $H^1(X_0(N), \mathbb{C}) \cong H^0(X_0(N), \Omega^1_{X_0(N)}) \oplus \overline{H^0(X_0(N), \Omega^1_{X_0(N)})}$ is compatible with the action of Hecke operators. It follows that the differential form $\omega = i^*\pi^*\omega_E$ determines a class in singular cohomology which is an eigenvector for all Hecke operators T_p , $p \nmid N$. We therefore make the following revised definition.

Definition 2.8. Let E be an elliptic curve over K of conductor \mathfrak{n} with $\operatorname{End}_K(E) = \mathbb{Z}^2$. We say that E is modular if there exists a non-zero class $c_E \in H^*(Y_0(\mathfrak{n}), \mathbb{C})$ such that for all places v not dividing \mathfrak{n} , $T_v c_E = a_v c_E$ (where $a_v = q_v + 1 - \#E(k(v))$).

Theorem 2.9. When $K = \mathbb{Q}$, Definition 2.8 is equivalent to Definition 2.2.

Proof (sketch). We note that the condition $\operatorname{End}_K(E) = \mathbb{Z}$ is automatic for an elliptic curve defined over $K = \mathbb{Q}$. We have seen that the Eichler-Shimura relation implies that our previous definition implies this one. In the other direction, one can use the Eichler-Shimura relation to show that if E is modular in this new sense, then there exists a surjective homomorphism $\pi': J_0(N) \to E'$, where E' is an elliptic curve of conductor dividing N such that for all primes p not dividing N, $a_p(E) = a_p(E')$. Faltings' theorem implies that there is an isogeny $f: E' \to E$, and we can take $\pi = f \circ \pi'$.

For this definition to make sense, we need to define the Hecke operators T_v . The definition is the same as before: for any place $v \nmid \mathfrak{n}$, there is a diagram



where the maps have finite fibres. We define the operator T_v on $H^*(Y_0(\mathfrak{n}),\mathbb{C})$ by the formula $T_v = \pi_{2,*} \circ \pi_1^*$.

Conjecture 2.10. Let E be an elliptic curve over the number field K, and suppose that $\operatorname{End}_K(E) = \mathbb{Z}$. Then E is modular.

Just as in the case $K = \mathbb{Q}$, the modularity of an elliptic curve E over K has consequences for its L-function.

Theorem 2.11. Let E be a modular elliptic curve of conductor \mathfrak{n} . Then there exists an L-function $L(\pi, s) = \prod_v L_v(\pi_v, s)$ such that:

²The elliptic curves with complex multiplication defined over K have to be treated separately; this is related to the fact that their ℓ -adic representations are reducible.

- 1. For all places $v \nmid \mathfrak{n}$, $L_v(\pi_v, s) = L_v(E, s)$.
- 2. $L(\pi, s)$ converges absolutely in a right half-plane, has an analytic continuation to the complex plane, and satisfies a functional equation there.

Proof (sketch). The cohomology class $c_E \in H^*(Y_0(\mathfrak{n}), \mathbb{C})$ determines a function $f: \mathrm{GL}_2(K) \backslash \mathrm{GL}_2(\mathbb{A}_K) \to \mathbb{C}$, an automorphic form which generates an automorphic representation π of $\mathrm{GL}_2(\mathbb{A}_K)$. Applying a generalized Mellin transform to f, we obtain a completed L-function $\Lambda(\pi,s)$ which has an Euler product which agrees with $\Lambda(E,s)$ at all $v \nmid \mathfrak{n}$, and which can be proved to have an analytic continuation to \mathbb{C} and a functional equation.

[We note that one can state a more refined version of the modularity conjecture under which we have $L(E,s) = L(\pi,s)$. We sketch this now. Let v be a finite place of K. The local Langlands conjecture for $GL_2(K_v)$, now a theorem, gives a bijection between two sets:

- The set of isomorphism classes of irreducible admissible $\mathbb{C}[\operatorname{GL}_2(K_v)]$ -modules π_v .
- The set of 2-dimensional Frobenius-semisimple Weil-Deligne representations (r_v, N_v) .

If π_v is an irreducible admissible $\mathbb{C}[\operatorname{GL}_2(K_v)]$ -module, then we write $\operatorname{rec}(\pi_v)$ for the associated Weil-Deligne representation. One knows how to associate to the elliptic curve E for every finite place v of K a Weil-Deligne representation (r_v, N_v) , which is Frobenius-semisimple. This allows us to associate to every finite place v of K the irreducible admissible representation $\pi_v(E)$ of $\operatorname{GL}_2(K_v)$ such that $\operatorname{rec}(\pi_v) = (r_v, N_v)$. The refined modularity conjecture is as follows:

Conjecture 2.12. Let $\mathcal{A} = \varinjlim_{U} H^{*}(Y_{U}, \mathbb{C})$, an admissible $\mathbb{C}[\operatorname{GL}_{2}(\mathbb{A}_{K}^{\infty})]$ -module, and let E be an elliptic curve over K with $\operatorname{End}_{K}(E) = \mathbb{Z}$. Then the representation $\pi(E) = \bigotimes'_{v} \pi_{v}(E)$ is a subquotient of \mathcal{A} .

It is an exercise in the representation theory of $GL_2(K_v)$ to show that this conjecture implies the previous one. One can show that if E satisfies the conclusion of this conjecture, then indeed L(E,s) admits an analytic continuation to $\mathbb C$ and satisfies the expected functional equation.]

What is known about the modularity conjecture? The best-studied case is when K is totally real. For example, we have the following result.

Theorem 2.13. Let K be a totally real field satisfying one of the following conditions:

- 1. $K = \mathbb{Q}$.
- 2. K/\mathbb{Q} is quadratic.
- 3. There exists a prime p such that K/\mathbb{Q} is a cyclic p-power extension unramified outside p.

Then every elliptic curve E over K is modular.

Very little is known beyond this. The case where K is a CM field (i.e. a totally imaginary extension of a totally real field) is a topic of current research. Beyond this case, we know essentially nothing.

We end with an important observation: the condition for E to be modular depends only on its ℓ -adic Galois representations $\rho_{E,\ell}: G_K \to \operatorname{GL}_2(\overline{\mathbb{Q}}_\ell)$. (We use the following notation: $G_K = \operatorname{Gal}(\overline{K}/K)$ is the absolute Galois group of K with respect to a fixed choice of algebraic closure. If v if a place of K, then an embedding $\overline{K} \to \overline{K}_v$ determines a choice of decomposition group $G_{K_v} \to G_K$.) In fact, there is no reason to restrict to elliptic curves in the definition of modularity:

Definition 2.14. Let ℓ be a prime, and let $\rho_{\ell}: G_K \to \operatorname{GL}_2(\overline{\mathbb{Q}}_{\ell})$ be a continuous, irreducible representation which is unramified at all but finitely many places of K. We say that ρ_{ℓ} is modular of weight 2 if there exists a non-zero ideal $\mathfrak{n} \subset \mathcal{O}_K$ and a non-zero class $c_{\rho_{\ell}} \in H^*(Y_1(\mathfrak{n}), \overline{\mathbb{Q}}_{\ell})$ such that for all places v not dividing \mathfrak{n} , $\rho_{\ell}|_{G_{K_v}}$ is unramified and $T_v c_E = (\operatorname{tr} \rho_{\ell}(\operatorname{Frob}_v))c_E$.

This wider context is essential for actually proving positive modularity results. In this definition, we set $Y_1(\mathfrak{n}) = Y_{U_1(\mathfrak{n})}$, where

$$U_1(\mathfrak{n}) = \left\{ \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in \operatorname{GL}_2(\widehat{\mathcal{O}}_K) \mid c \equiv 0 \bmod \mathfrak{n}, d \equiv 1 \bmod \mathfrak{n} \right\}.$$

Conjecture 2.15. Let ℓ be a prime, and let $\rho_{\ell}: G_K \to \operatorname{GL}_2(\overline{\mathbb{Q}}_{\ell})$ be a continuous, irreducible representation which is unramified at all but finitely many places of K. Suppose that for each place $v|\ell$ of K, $\rho_{\ell}|_{G_{K_v}}$ is of weight 2. Then ρ_{ℓ} is modular of weight 2.

Unfortunately we do not have time here to explain what it means for $\rho_{\ell}|_{G_{K_v}}$ to be of weight 2 when $v|\ell$; it is essentially a condition in p-adic Hodge theory, which is satisfied by the Galois representations attached to elliptic curves. The reason for switching to $Y_1(\mathfrak{n})$ in the statement of Conjecture 2.15 is that the Hecke eigenvalue systems appearing in $H^*(Y_0(\mathfrak{n}),\mathbb{C})$ are supposed to be attached to Galois representations with determinant of the form $\epsilon_\ell \psi$, where ϵ_ℓ is the ℓ -adic cyclotomic character and $\psi: G_K \to \overline{\mathbb{Q}}_\ell^\times$ is an everywhere unramified character of finite order (in other words, a character of the narrow ideal class group of K). In order to allow Galois representations without restriction on the determinant, we must pass to $H^*(Y_1(\mathfrak{n}),\mathbb{C})$. The reasons for this are once more tied up with the local Langlands conjectures for GL_2 and the representation theory of the group $\mathrm{GL}_2(K_v)$.

References

The cohomology of the spaces Y_U is an important topic in the theory of automorphic forms, in particular because it is possible, in principle, to program a computer to compute the cohomology groups $H^*(Y_U, \mathbb{C})$, together with the action of the Hecke operators. See for example [Cre84], which contains computations in the case where K is an imaginary quadratic field, or [RcS13] for a more recent reference. For a concise statement of the modularity conjecture, together with its extension to abelian varieties of GL_2 -type, see Taylor's ICM article [Tay95], which also includes the definition of what it means for a Galois representation to be of weight 2.

For an excellent introduction to the representation theory of $\operatorname{GL}_2(K_v)$, where v is a finite place of the number field K, see [BH06]. The properties of L-functions of automorphic representations of $\operatorname{GL}_2(\mathbb{A}_K)$ were established in the landmark work of Jacquet–Langlands [JL70]. The proof of the analytic continuation and functional equation of $L(\pi, s)$ is a far-reaching generalization of the technique of applied to a classical holomorphic modular form in the first lecture. A more approachable reference, which treats the same material, is the book of Gelbart [Gel75].

The modularity of elliptic curves over real quadratic fields is due to Freitas, Le Hung, and Siksek [FLHS15]. The modularity of elliptic curves over real abelian p-extensions ramified only at p is proved in [Tho].

2.3 Lecture 3

We now know what it means for an elliptic curve over an arbitrary number field to be modular, and that this implies that the L-function of the corresponding elliptic curve has all desired properties. However, this is only useful if elliptic curves can be proved to be modular!

It is often possible to prove that a given elliptic curve is modular by using computer calculation, provided one is a context where the Galois representations attached to automorphic forms have been proved to exist. Indeed, let K be a number field, let ℓ be a prime, and let S be a finite set of finite places of K, containing the non-archimedean ones. The Faltings–Serre method gives an effective constant $C(K, S, \ell)$ satisfying the following condition: let $\rho, \rho' : G_K \to \mathrm{GL}_2(\mathbb{Q}_\ell)$ be two continuous semisimple representations which are unramified outside S, and suppose that that $\mathrm{tr}\,\rho(\mathrm{Frob}_v) = \mathrm{tr}\,\rho'(\mathrm{Frob}_v)$ for all finite places v of K such that $v \notin S$ and $q_v < C$. Then $\rho \cong \rho'$.

This can be used to prove the modularity of a given elliptic curve E over K, provided one knows that for any Hecke eigenclass $c \in H^*(Y_0(\mathfrak{n}), \mathbb{C})$ and prime ℓ , there exists a continuous representation $\rho_{c,\ell}: G_K \to \mathrm{GL}_2(\overline{\mathbb{Q}}_\ell)$ unramified outside $\mathfrak{n}\ell$ such that for each finite place $v \nmid \mathfrak{n}\ell$ of K, tr $\rho_{c,\ell}(\mathrm{Frob}_v)$ is equal to the eigenvalue of T_v on c. Indeed, one then need only find a suitable eigenclass c_E and check agreement of sufficiently many Frobenius traces with Hecke eigenvalues. For example, the elliptic curve

$$y^{2} + xy = x^{3} + \frac{3 + \sqrt{-3}}{2}x^{2} + \frac{1 + \sqrt{-3}}{2}x$$

over $K = \mathbb{Q}(\sqrt{-3})$ has been proved to be modular by applying this technique. The repersentations $\rho_{c,\ell}$ have been proved to exist for any (totally real or totally imaginary) CM field K [HLTT16].

However, it is desirable to have more general results. To this end, the notion of potential modularity was introduced by Taylor. Its effectiveness is based on a fundamental property of the *L*-functions attached to Galois representations, namely their compatibility with induction, which we refer to as "Artin formalism" (the original reference is [Art24], which is also the paper where Artin defined the *L*-function of a non-abelian Galois representation for the first time).

Proposition 2.16. Let $\iota : \overline{\mathbb{Q}}_{\ell} \cong \mathbb{C}$ be an isomorphism. Let E/K be an extension of number fields, and let $\rho_{\ell} : G_E \to \operatorname{GL}_2(\overline{\mathbb{Q}}_{\ell})$ be a continuous representation which is unramified at all but finitely places of K. Then $L(\iota \rho_{\ell}, s) = L(\iota \operatorname{Ind}_{G_E}^{G_K} \rho_{\ell}, s)$.

Proof. It suffices to show the equality one Euler factor at a time, i.e. for every place v of K,

$$L_v(\operatorname{Ind}_{G_E}^{G_K} \rho_{\ell}, s) = \prod_{w|v} L_w(\rho_{\ell}, s),$$

the product running over the set of places w of E lying above K. By Mackey's formula, we have

$$\operatorname{Res}_{G_{K_v}}^{G_K}\operatorname{Ind}_{G_E}^{G_K}\rho_{\ell}\cong \oplus_{w|v}\operatorname{Ind}_{G_{E_w}}^{G_{K_v}}\operatorname{Res}_{G_{E_w}}^{G_E}\rho_{\ell}.$$

It is therefore enough to show that for each any place w of E lying above the place v of K, we have

$$L_v(\operatorname{Ind}_{G_{E_w}}^{G_{K_v}} \rho_{\ell}|_{G_{E_w}}) = L_w(\rho_{\ell}|_{G_{E_w}}).$$

Let us just treat the case where $\rho_{\ell}|_{G_{E_w}}$ is unramified and E_w/K_v is unramified, leaving the general case to the exercises. (This proves the result already at all but finitely many places of K.) In this case we can reformulate the result as follows: let $\Gamma = \mathbb{Z}$, with generator ϕ , and let $\Delta = n\mathbb{Z}$ for some $n \neq 0$. Let $\chi : \Delta \to \mathbb{C}^{\times}$ be a character. Then we must show that

$$\det(1 - X\phi \mid \operatorname{Ind}_{\Gamma}^{\Delta} \chi) = \det(1 - X^n \phi^n \mid \chi).$$

This can be checked directly.

³We are brushing under the rug here the issue of defining the L-factors of $\rho_{\ell}|_{G_{K_v}}$ at the ℓ -adic places v of K. This is possible when ρ_{ℓ} is assumed to be de Rham, in the sense of p-adic Hodge theory. Indeed, in this case, Fontaine showed how to attach a Weil–Deligne representation to $\rho_{\ell}|_{G_{K_v}}$, and one should define the L-factor of ρ_{ℓ} at v to be the L-factor of this Weil–Deligne representation. One can avoid such difficult results by assuming that ρ_{ℓ} lives in a compatible system of Galois representations, for example those arising from an elliptic curve.

Definition 2.17. Let K be a number field and let ℓ be a prime. Let $\rho_{\ell}: G_K \to \operatorname{GL}_2(\overline{\mathbb{Q}}_{\ell})$ be a continuous irreducible representation. We say that ρ_{ℓ} is potentially modular of weight 2 if there exists a finite Galois extension E/K such that $\rho_{\ell}|_{G_E}$ is irreducible and modular of weight 2.

Theorem 2.18. Suppose that ρ_{ℓ} is potentially modular of weight 2. Then $L(\iota \rho_{\ell}, s)$ admits a meromorphic continuation to the whole complex plane and satisfies a functional equation.

Note that if ρ_{ℓ} arises from an elliptic curve over K, then the meromorphic continuation of $L(\iota \rho_{\ell}, s)$ is enough to be able to formulate BSD unconditionally!

Proof (Sketch). Let $G = \operatorname{Gal}(E/K)$. By Brauer's theorem, we can find soluble subgroups $H_i \subset G$, integers $n_i \in \mathbb{Z}$, and characters $\chi_i : H_i \to \overline{\mathbb{Q}}_{\ell}^{\times}$ such that $\mathbf{1} = \sum_i n_i \operatorname{Ind}_{H_i}^G \chi_i$ (identity in the Grothendieck group of representations of G). Let $K_i = E^{H_i}$. Taking the tensor product with ρ_{ℓ} , we get an identity

$$\rho_{\ell} = \sum_{i} n_{i} \operatorname{Ind}_{G_{K_{i}}}^{G_{K}} (\chi_{i} \otimes \rho_{\ell}|_{G_{K_{i}}}).$$

This implies a corresponding identity of L-functions

$$L(\iota \rho_{\ell}, s) = \prod_{i} L(\iota \operatorname{Ind}_{G_{K_{i}}}^{G_{K}} \rho_{\ell}|_{G_{K_{i}}} \otimes \chi_{i}, s)^{n_{i}} = \prod_{i} L(\iota \rho_{\ell}|_{G_{K_{i}}} \otimes \chi_{i}, s)^{n_{i}}.$$

We now need to use two critical pieces of information. The first is that modularity can be descended along a soluble extension of number fields. (This is a hard theorem which is due in general to Langlands for GL_2 [Lan80]; the proof is reduced by induction to the case of a cyclic extension.) The second is that modularity is preserved under character twist. (This is much easier.) It follows that each of the Galois representations $\rho_{\ell}|_{G_{K_i}} \otimes \chi_i$ is modular, and hence that each of the *L*-functions $L(\iota \rho_{\ell}|_{G_{K_i}} \otimes \chi_i, s)$ has an analytic continuation to $\mathbb C$ and satisfies a functional equation relating s and 1 - s. It follows that 1 - s has a meromorphic continuation to 1 - s (some of the 1 - s has a meromorphic and satisfies a functional equation relating 1 - s has a meromorphic continuation to 1 - s has a meromorphic cont

Theorem 2.19 (Taylor). Let E be an elliptic curve over a totally real field. Then E is potentially modular. Consequently, the L-function L(E,s) admits a meromorphic continuation to \mathbb{C} and satisfies the expected functional equation.

Another application of potential modularity is the proof of the Sato-Tate conjecture. We recall that if E is an elliptic curve over a number field K without complex multiplication, then the Sato-Tate conjecture for E predicts that the quantities $a_v/2\sqrt{q_v} \in [-1,1]$ are equidistributed with respect to the Sato-Tate measure

$$\frac{2}{\pi}\sqrt{1-t^2}\,dt$$

as v varies through all finite places of the number field K at which E has good reduction. Serre had observed already in the 60's [Ser98, Ch. I, Appendix] that this conjecture would follow if one could show that for each $n \ge 1$, the symmetric power L-function

$$L(\iota \operatorname{Sym}^n \rho_{E,\ell}, s)$$

had a meromorphic continuation to \mathbb{C} which was holomorphic and non-vanishing on the line Re s=1+n/2. This property is known for the *L*-functions of modular Galois representations (once we have defined what it means for an (n+1)-dimensional Galois representation to be modular). In fact, the same argument as above shows that it is enough to prove that the symmetric power Galois representations are merely potentially modular. In this way, one can prove the following theorem:

Theorem 2.20 (Clozel-Harris-Shepherd-Barron-Taylor-Barnet-Lamb-Gee-Geraghty). Let E be an elliptic curve over a totally real field K without complex multiplication. Then the Sato-Tate conjecture holds for E.

References

For a recent paper illustrating the Faltings–Serre method, see [DGP10]. The technique of potential modularity was introduced in the paper [Tay02]. It is very flexible and powerful, and can be applied to much more general 2-dimensional Galois representations over totally real fields than ones arising from elliptic curves. See [BLGGT14] for the most general results currently available (and also results for some Galois representations in dimension $n \geq 2$).

The Sato-Tate conjecture can be formulated for any motive (or indeed, for its attached Galois representations). See [Ser12, Ch. 8] for a nice discussion of this. The general conjecture can also be shown to follow from properties of *L*-functions, but establishing these in general seems an impossibly hard problem. The best results available for 2-dimensional Galois representations can be found in [BLGG11].

2.4 Lecture 4

So far we have discussed only what it means for 2-dimensional Galois representations of weight 2 to be modular. However, this is clearly not the only case of interest! For example, we may be interested in:

- Galois representations $\operatorname{Sym}^n \rho_{E,\ell}: G_{\mathbb{Q}} \to \operatorname{GL}_{n+1}(\mathbb{Q}_{\ell})$, where E is an elliptic curve over \mathbb{Q} . We have seen that these are relevant for the Sato-Tate conjecture.
- Galois representations $\rho_{A,\ell}: G_{\mathbb{Q}} \to \mathrm{GSp}_{2g}(\mathbb{Q}_{\ell})$, where A is an abelian variety over \mathbb{Q} of dimension g > 1. These include the Galois representations attached to hyperelliptic curves of genus g.

In this lecture we aim to put these questions in a more general context, namely that of the Langlands program. The punchline is that using ideas of the Langlands program, it is possible to make a completely precise conjecture which generalizes the Shimura–Taniyama–Weil conjecture to abelian varieties of higher dimension, and which can be tested by computer (see [BK14, Gro16]). In order to explain where this comes from, we have to introduce a number of new concepts.

Let K be a number field, and let G be a connected reductive group over K. For reasons of simplicity of exposition, we are going to assume that G is split (i.e. that G contains a split maximal torus), although everything we say here generalizes to the case of an arbitrary reductive group. For example, we could take G to be one of the following:

- $G = GL_n$.
- $G = \operatorname{Sp}_{2n}$.
- $G = SO_n$, the split orthogonal group defined by the symmetric bilinear form $\langle x, y \rangle = \sum_{i=1}^n x_i y_{n+1}$ on K^n .

Let $Z \subset G$ denote the centre of G, and $\omega : Z(K) \setminus Z(\mathbb{A}_K) \to \mathbb{C}^{\times}$ be a continuous character. The quotient $G(K)Z(\mathbb{A}_K) \setminus G(\mathbb{A}_K)$ has finite volume, and we can define $L^2_{\omega,0}(G(K) \setminus G(\mathbb{A}_K))$ to be the Hilbert space of functions $f : G(K) \setminus G(\mathbb{A}_K) \to \mathbb{C}$ satisfying the following conditions:

- f has central character ω : for all $g \in G(\mathbb{A}_K)$, $z \in Z(\mathbb{A}_K)$, $\gamma \in G(K)$, we have $f(\gamma zg) = \omega(z)f(g)$.
- f is cuspidal: for all proper parabolic subgroups $P \subset G$ of unipotent radical N, and for all $g \in G(\mathbb{A}_K)$, we have

$$\int_{n\in N(K)\setminus N(\mathbb{A}_K)} f(ng) \, dn = 0.$$

Then $L^2_{\omega,0}(G(K)\backslash G(\mathbb{A}_K))$ is a continuous representation of the locally compact topological group $G(\mathbb{A}_K)$, which acts by right translation. It is a unitary representation if the character ω is unitary.

Definition 2.21. A cuspidal automorphic representation of $G(\mathbb{A}_K)$ is a closed irreducible subrepresentation $\pi \subset L^2_{\omega,0}(G(K)\backslash G(\mathbb{A}_K))$ (for some choice of ω).

Conjecturally, automorphic representations of the group $G(\mathbb{A}_K)$ are related to Galois representations valued in another group \widehat{G} , the so-called dual group. The dual group \widehat{G} is a connected reductive group over \mathbb{C} which is defined in terms of G using the classification of reductive groups in terms of roots and weights. It is a simple matter to write it down in any of the above cases. For example, in the above examples we have

- If $G = GL_n$, then $\widehat{G} = GL_n$.
- If $G = \operatorname{Sp}_{2n}$, then $\widehat{G} = \operatorname{SO}_{2n+1}$.
- If $G = SO_{2n+1}$, then $\widehat{G} = Sp_{2n}$.

Conjecture 2.22. Let ℓ be a prime, and fix an isomorphism $\iota : \overline{\mathbb{Q}}_{\ell} \cong \mathbb{C}$. Then there is a correspondence

$$\left\{ \begin{array}{c} \rho: G_K \to \widehat{G}(\overline{\mathbb{Q}}_\ell) \\ algebraic, \ irreducible \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} \pi \ algebraic \ cuspidal \ automorphic \\ representation \ of \ G(\mathbb{A}_K) \end{array} \right\}.$$

In order for this conjecture to make sense, we need to explain some of the terms.⁴ A Galois representation $\rho: G_K \to \widehat{G}(\overline{\mathbb{Q}}_{\ell})$ is said to be algebraic if it satisfies the following conditions:

- ρ is continuous and is unramified at all but finitely many primes.
- For each place $v|\ell$ of K, $\rho|_{G_{K_v}}$ is de Rham in the sense of p-adic Hodge theory.

It is irreducible if the image of ρ is contained in no proper parabolic subgroup of \widehat{G} ; if $\widehat{G} = GL_n$, this is equivalent to the usual notion of irreducibility.

A cuspidal automorphic representation π of $G(\mathbb{A}_K)$ admits a factorization $\pi = \otimes'_{v \in \mathcal{M}_K} \pi_v$ as a restricted tensor product, where each π_v is an irreducible representation of $G(K_v)$. For all but finitely many finite places v, the unramified local Langlands correspondence attaches to π_v an unramified homomorphism ϕ_{π_v} : $W_{K_v} \to \widehat{G}(\mathbb{C})$. If v is an archimedean place, then the local Langlands correspondence for $G(K_v)$ associates to π_v a continuous homomorphism $\phi_{\pi_v}: W_{K_v} \to \widehat{G}(\mathbb{C})$. If $K_v = \mathbb{C}$, then $W_{K_v} = \mathbb{C}^{\times}$. If $K_v = \mathbb{R}$, then $W_{K_v} = \mathbb{C}^{\times} \cup j\mathbb{C}^{\times}$, otherwise known as the units in the Hamiltonian quaternions. In either case we can consider the restriction $\phi_{\pi_v}|_{\mathbb{C}^{\times}}$, and we say that π_v is algebraic if it satisfies the following condition:

• The restriction $\phi_{\pi_v}|_{\mathbb{C}^\times}: \mathbb{C}^\times \to \widehat{G}(\mathbb{C})$ arises from a homomorphism $\mathbb{G}_m \to \widehat{G}$ of algebraic groups over \mathbb{C} .

We say that π is algebraic if π_v is algebraic for each place $v \mid \infty$ of K.

If ρ and π are related under the correspondence of Conjecture 2.22, then for all but finitely many finite places v, $\iota\rho|_{W_{K_v}}$ should be $\widehat{G}(\mathbb{C})$ -conjugate to ϕ_{π_v} .

⁴To get a conjecture that is close to being true, we should also ask that π is everywhere tempered. We avoid going into further details here.

Even with the above desiderata in place, the above conjecture is so imprecise that it hardly deserves to be called a conjecture. For example, what is a 'correspondence'? When $G = GL_n$, we can be more precise:

Conjecture 2.23. Let ℓ be a prime, and fix an isomorphism $\iota : \overline{\mathbb{Q}}_{\ell} \cong \mathbb{C}$. Then there is a bijection

$$\left\{\begin{array}{c} \rho: G_K \to \operatorname{GL}_n(\overline{\mathbb{Q}}_\ell) \\ algebraic, \ irreducible \end{array}\right\} \leftrightarrow \left\{\begin{array}{c} \pi \ \ cuspidal \ \ algebraic \ \ automorphic \\ representation \ \ of \ GL_n(\mathbb{A}_K) \end{array}\right\},$$

which is uniquely characterized by the following property: if ρ and π are related under this correspondence, then for all but finitely many finite places v, $\iota\rho|_{W_{K_v}}$ and ϕ_{π_v} are $\mathrm{GL}_n(\mathbb{C})$ -conjugate.

In light of this more precise conjecture, one may ask why it is worthwhile to consider automorphic representations of groups other than GL_n . For example, suppose that A is an abelian variety of dimension g over \mathbb{Q} and ℓ is a prime such that the associated ℓ -adic representation $\rho_{A,\ell}$ has image equal to $GSp_{2g}(\mathbb{Z}_{\ell})$. We see that $\rho_{A,\ell}$ should determine automorphic representations both of the group SO_{2g+1} and of the group GL_{2g} . What is the benefit of considering SO_{2g+1} ?

The point is that, depending on the behaviour of the infinite component π_{∞} , we may be able to access the finite part π^{∞} in other ways. In order to simplify the discussion, we now assume that G is semisimple (i.e. that its centre is finite; this is the case if $G = \operatorname{Sp}_{2n}$ or $G = \operatorname{SO}_n$). If $U \subset G(\mathbb{A}_K^{\infty})$ is an open compact subgroup, we define a space

$$Y_U = G(K) \backslash G(\mathbb{A}_K^{\infty}) \times X/U,$$

where $X = G(K \otimes_{\mathbb{Q}} \mathbb{R})/U_{\infty}$ and U_{∞} is a maximal compact subgroup of $G(K \otimes_{\mathbb{Q}} \mathbb{R})$. The space Y_U is what we call an arithmetic locally symmetric space; it is a disjoint union of finitely many quotients of X, a Riemannian symmetric space, by arithmetic subgroups $\Gamma \subset G(K)$. This is the analogue for a general group G of the space Y_U defined in the second lecture for GL_2 . In general there is an injection

$$\bigoplus_{\pi} (\pi^{\infty})^U \otimes_{\mathbb{C}} H^*(\mathfrak{g}, U_{\infty}; \pi_{\infty}) \hookrightarrow H^*(Y_U, \mathbb{C}),$$

where $\mathfrak{g} = (\operatorname{Lie} G(K \otimes_{\mathbb{Q}} \mathbb{R})) \otimes_{\mathbb{R}} \mathbb{C}$ and $H^*(\mathfrak{g}, U_{\infty}; \pi_{\infty})$ is the (finite-dimensional) so-called $(\mathfrak{g}, U_{\infty})$ -cohomology of π_{∞} , and the sum runs over the set of cuspidal automorphic representations π of $G(\mathbb{A}_K)$. Thus the automorphic representations π such that π_{∞} has non-trivial $(\mathfrak{g}, U_{\infty})$ -cohomology can be studied through the (singular) cohomology of the spaces Y_U . If $G = \operatorname{GL}_2$ then the automorphic representations which should correspond to elliptic curves are among these, and the automorphic representations which contribute are uniquely determined by their corresponding systems of Hecke eigenvalues, which is why we were able to define modularity of elliptic curves in the way we did.

If A is an abelian variety over \mathbb{Q} of dimension g > 1 with $\operatorname{End}_{\mathbb{Q}}(A) = \mathbb{Z}$, then the representations $\rho_{A,\ell} : G_{\mathbb{Q}} \to \operatorname{GSp}_{2g}(\mathbb{Q}_{\ell})$ should give rise to an algebraic cuspidal automorphic representation π_A of $\operatorname{GL}_{2g}(\mathbb{A}_{\mathbb{Q}})$. However, the infinite component $\pi_{A,\infty}$ should have vanishing $(\mathfrak{g}, U_{\infty})$ -cohomology, so we cannot define modularity of an abelian variety in the same way as we did for elliptic curves by using an explicit realization inside the singular cohomology of an arithmetic locally symmetric space.

If A is an abelian variety of dimension g=2 and $\operatorname{End}_{\mathbb{Q}}(A)=\mathbb{Z}$, then calculations due to Gross using a more refined version of the Langlands conjectures [Gro16] imply that the representations $\rho_{A,\ell}:G_{\mathbb{Q}}\to\operatorname{GSp}_4(\mathbb{Q}_\ell)$ should correspond to automorphic representations π of $G(\mathbb{A}_{\mathbb{Q}})=\operatorname{SO}_5(\mathbb{A}_{\mathbb{Q}})$ which can be realized inside spaces

⁵In the best possible situation, we hope to divide each side of the correspondence into 'packets'. The packets should be in bijective correspondence, and we hope to be able to describe the packets explicitly. The conjecture for GL_n is simpler because in this case all the packets should be singletons.

⁶We note in passing that the group G with dual group GSp_{2g} is the general spin group $\operatorname{GSpin}_{2g+1}$, which is an extension of SO_{2g+1} by \mathbb{G}_m . However, the automorphic representations of $\operatorname{GSpin}_{2g+1}(\mathbb{A}_{\mathbb{Q}})$ corresponding to abelian varieties should have character twists which descend to automorphic representations by SO_{2g+1} . Compare [Gro16, §5]. This is already the case when g=1: we can associate to an elliptic curve E over \mathbb{Q} an automorphic representation of GL_2 of trivial central character, which therefore descends to $\operatorname{PGL}_2 \cong \operatorname{SO}_3$.

of holomorphic modular forms on the spaces Y_U , which admit a complex structure. Under the exceptional isomorphism $SO_5 \cong PSp_4$, these correspond to Siegel modular forms of genus 2 and weight 2, which are exactly the forms appearing in the conjectures of Brumer and Kramer [BK14], and which have been the subject of computer calculation by Poor and Yuen [PY15]. It follows that in order to study these Galois representations $\rho_{A,\ell}$ by computer calculation, it is necessary to make use of the framework of the Langlands program on the intermediate group SO_{2g+1} ! We refer the reader again to [Gro16] for a beautiful exposition of a precise generalization of the Shimura–Weil–Taniyama conjecture to abelian varieties of arbitrary dimension g > 1.

References

An excellent reference for the basic theory of automorphic representations is the Corvallis conference proceedings [BC79]. In particular, we mention the articles of Springer (which gives an introduction to the theory of reductive groups), Flath (which describes the decomposition of a representation of $G(\mathbb{A}_K)$ as a restricted tensor product) and Borel (which describes the dual group of a reductive group, as well as a large part of the Langlands conjectures in the local case). Conjecture 2.23 was first stated in [Clo90]. The reciprocity conjecture for a general reductive group was first stated by Buzzard and Gee [BG14], although special cases were considered earlier by Gross (see e.g. [Gro99]).

References

- [Art24] E. Artin. über eine neue art von L-Reihen. Abh. Math. Sem. Univ. Hamburg, 3(1):89–108, 1924.
- [BC79] Armand Borel and W. Casselman, editors. Automorphic forms, representations and L-functions. Part 1, Proceedings of Symposia in Pure Mathematics, XXXIII. American Mathematical Society, Providence, R.I., 1979.
- [BCDT01] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor. On the modularity of elliptic curves over **Q**: wild 3-adic exercises. *J. Amer. Math. Soc.*, 14(4):843–939, 2001.
- [BG14] Kevin Buzzard and Toby Gee. The conjectural connections between automorphic representations and Galois representations. In *Automorphic forms and Galois representations*. Vol. 1, volume 414 of *London Math. Soc. Lecture Note Ser.*, pages 135–187. Cambridge Univ. Press, Cambridge, 2014.
- [BH06] Colin J. Bushnell and Guy Henniart. The local Langlands conjecture for GL(2), volume 335 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2006.
- [BK14] Armand Brumer and Kenneth Kramer. Paramodular abelian varieties of odd conductor. *Trans. Amer. Math. Soc.*, 366(5):2463–2516, 2014.
- [BLGG11] Thomas Barnet-Lamb, Toby Gee, and David Geraghty. The Sato-Tate conjecture for Hilbert modular forms. J. Amer. Math. Soc., 24(2):411–469, 2011.
- [BLGGT14] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor. Potential automorphy and change of weight. *Ann. of Math.* (2), 179(2):501–609, 2014.
- [Clo90] Laurent Clozel. Motifs et formes automorphes: applications du principe de fonctorialité. In Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), volume 10 of Perspect. Math., pages 77–159. Academic Press, Boston, MA, 1990.

- [Cre84] J. E. Cremona. Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields. *Compositio Math.*, 51(3):275–324, 1984.
- [DDT94] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. In *Current developments in mathematics*, 1995 (Cambridge, MA), pages 1–154. Int. Press, Cambridge, MA, 1994.
- [DGP10] Luis Dieulefait, Lucio Guerberoff, and Ariel Pacetti. Proving modularity for a given elliptic curve over an imaginary quadratic field. *Math. Comp.*, 79(270):1145–1170, 2010.
- [DR73] P. Deligne and M. Rapoport. Les schémas de modules de courbes elliptiques. pages 143–316. Lecture Notes in Math., Vol. 349, 1973.
- [DS05] Fred Diamond and Jerry Shurman. A first course in modular forms, volume 228 of Graduate Texts in Mathematics. Springer-Verlag, New York, 2005.
- [FLHS15] Nuno Freitas, Bao V. Le Hung, and Samir Siksek. Elliptic curves over real quadratic fields are modular. *Invent. Math.*, 201(1):159–206, 2015.
- [Gel75] Stephen S. Gelbart. Automorphic forms on adèle groups. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1975. Annals of Mathematics Studies, No. 83.
- [Gro99] Benedict H. Gross. Algebraic modular forms. Israel J. Math., 113:61–93, 1999.
- [Gro16] B. Kh. Gross. On the Langlands correspondence for symplectic motives. *Izv. Ross. Akad. Nauk Ser. Mat.*, 80(4):49–64, 2016.
- [HLTT16] Michael Harris, Kai-Wen Lan, Richard Taylor, and Jack Thorne. On the rigid cohomology of certain Shimura varieties. *Res. Math. Sci.*, 3:Paper No. 37, 308, 2016.
- [JL70] H. Jacquet and R. P. Langlands. *Automorphic forms on* GL(2). Lecture Notes in Mathematics, Vol. 114. Springer-Verlag, Berlin-New York, 1970.
- [Lan80] Robert P. Langlands. Base change for GL(2), volume 96 of Annals of Mathematics Studies. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980.
- [PY15] Cris Poor and David S. Yuen. Paramodular cusp forms. *Math. Comp.*, 84(293):1401–1438, 2015.
- [RcS13] Alexander D. Rahm and Mehmet Haluk , Sengün. On level one cuspidal Bianchi modular forms. LMS J. Comput. Math., 16:187–199, 2013.
- [Ser98] Jean-Pierre Serre. Abelian l-adic representations and elliptic curves, volume 7 of Research Notes in Mathematics. A K Peters, Ltd., Wellesley, MA, 1998. With the collaboration of Willem Kuyk and John Labute, Revised reprint of the 1968 original.
- [Ser12] Jean-Pierre Serre. Lectures on $N_X(p)$, volume 11 of Chapman & Hall/CRC Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2012.
- [Tay95] Richard Taylor. Representations of Galois groups associated to modular forms. In *Proceedings* of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), pages 435–442. Birkhäuser, Basel, 1995.
- [Tay02] Richard Taylor. Remarks on a conjecture of Fontaine and Mazur. J. Inst. Math. Jussieu, $1(1):125-143,\ 2002.$
- [Tho] Jack A. Thorne. Elliptic curves over \mathbb{Q}_{∞} are modular. To appear in J. Eur. Math. Soc.
- [TW95] Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math.* (2), 141(3):553–572, 1995.

[Wil95] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. Ann. of Math. (2), 141(3):443-551, 1995.

3 Exercises

3.1 Exercises to lecture 1

1. In this exercise we describe the complex structure on $X_0(N)$. We begin with a useful fact: let

$$\mathcal{D} = \{ \tau \in \mathbb{H} \mid |\operatorname{Re}(\tau) \le \frac{1}{2}, |\tau| \ge 1. \}$$

Then \mathcal{D} is a fundamental domain for the action of Γ on \mathbb{H} : any element of \mathbb{H} is conjugate to an element of \mathcal{D} , and if $\tau_1, \tau_2 \in \mathcal{D}$ are distinct points such that $\gamma \tau_1 = \tau_2$ for some $\gamma \in \Gamma$, then either τ_1, τ_2 both lie on the vertical boundary components and $\tau_1 = \tau_2 \pm 1$, or τ_1, τ_2 both lie on the horizontal boundary component and $\tau_1 = -1/\tau_2$.

- (a) Use this to show that Γ acts properly discontinuously on \mathbb{H} .
- (b) Show that Γ acts properly discontinuously on \mathbb{H}_{∞} . Deduce that $X_0(N)$ is a compact Hausdorff topological space.
- (c) Let $f: \mathbb{H}_{\infty} \to X_0(N)$ denote the tautological map. If $\tau \in \mathbb{H}$, show that $C_{\tau} = \operatorname{Stab}_{\Gamma_0(N)}(\tau)/\{\pm 1\}$ is cyclic of order $e_{\tau} \in \{1, 2, 3\}$. We define a complex chart around $f(\tau)$ as follows. Let $U_{\tau} \subset \mathbb{H}$ be an open neighbourhood stable under the action of C_{τ} , and such that for all $\gamma \in \Gamma_0(N)$, if $\gamma U_{\tau} \cap U_{\tau} \neq \emptyset$ then $\gamma \in C_{\tau}$. Let $m: \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ be a Möbius transformation which sends τ to 0 and $\overline{\tau}$ to ∞ . Then $mC_{\tau}m^{-1}$ acts by rotations around 0, and $z \mapsto m(z)^{e_{\tau}}$ is a chart on $C_{\tau} \setminus U_{\tau} \subset X_0(N)$.
- (d) We define a chart around $\infty \in \mathbb{H}_{\infty}$ as follows. Show that $C_{\infty} = \operatorname{Stab}_{\Gamma_0(N)}(\infty)/\{\pm 1\}$ is generated by the transformation $\tau \mapsto \tau + 1$. Show that we can find an open neighbourhood U_{∞} of ∞ in \mathbb{H}_{∞} , stable under $\tau \mapsto \tau + 1$, such that for all $\gamma \in \Gamma$, $\gamma U_{\infty} \cap U_{\infty} \neq \emptyset \Rightarrow \gamma \in C_{\infty}$. Let $q = e^{2\pi i \tau}$. Show that q extends to a homeomorphism from $C_{\infty} \setminus U_{\infty}$ to an open subset of the unit disc, which sends ∞ to 0. We take q to be a chart on $C_{\infty} \setminus U_{\infty}$. Explain how to extend this construction to give charts around each point of $X_0(N) Y_0(N)$.
- (e) Show that the above collection of charts makes $X_0(N)$ into a compact, connected Riemann surface.
- (f) (*) Show that if p is a prime such that $p \equiv -1 \mod 12$, then $\Gamma_0(p)/\{\pm 1\}$ contains no non-trivial elements of finite order. Apply the Riemann–Hurwitz theorem to the map $X_0(p) \to X_0(1)$ to calculate the genus of $X_0(p)$. (It may be helpful to note that the natural map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{F}_p)$ is surjective.)
- (g) (*) Visit the L-functions and modular forms database, and find out how many isogeny classes there are of elliptic curves over \mathbb{Q} of conductor 11. In light of the modularity theorem, how is this related to the previous exercise?
- 2. In this exercise we discuss Lemma 2.3. Suppose that E is an elliptic curve over \mathbb{Q} of conductor N, and that $\pi: J_0(N) \to E$ is a surjective homomorphism. Let p be a prime not dividing N.
 - (a) It is a fact that both $J_0(N)$ and E have good reduction at the prime p, so extend to abelian schemes \mathcal{J} and \mathcal{E} over \mathbb{Z}_p . Use the universal property of the Néron model to show that T_p and π extend to maps $\mathcal{J} \to \mathcal{J}$ and $\mathcal{J} \to \mathcal{E}$, respectively.
 - (b) Let $\phi_J \in \text{End}(\mathcal{J}_{\mathbb{F}_p})$, $\phi_E \in \text{End}(\mathcal{E}_{\mathbb{F}_p})$ be the Frobenius endomorphisms of the respective special fibres. The Eichler–Shimura relation says that

$$T_p \mod p = \phi_J + \widehat{\phi}_J$$

in $\operatorname{End}(\mathcal{J}_{\mathbb{F}_p})$, where hat denotes dual isogeny. (It is usually proved using the interpretation of $X_0(N)$ as a moduli space of elliptic curves; see e.g. [DS05, Ch. 8].) On the other hand, we have the relation

$$\phi_E + \widehat{\phi}_E = [a_p]$$

in $\operatorname{End}(\mathcal{E}_{\mathbb{F}_p})$. (This is easier.) Use these relations to show that $\pi \circ T_p = [a_p] \circ \pi$, as claimed in the lecture.

3.2 Exercises to lecture 2

- 1. If $K = \mathbb{Q}$, then a class in $H^1(X_0(N), \mathbb{C})$ with rational Hecke eigenvalues determines a modular elliptic curve. This need not be the case in general. If K is a number field and A is an abelian surface over K such that $\operatorname{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q} = B$ is a non-split quaternion algebra, we get for any prime ℓ a representation $\rho_{A,\ell}: G_K \to \operatorname{Aut}_{B\otimes_{\mathbb{Q}}\mathbb{Q}_{\ell}}(V_{\ell}(A)) \cong (B^{\operatorname{op}} \otimes_{\mathbb{Q}} \mathbb{Q}_{\ell})^{\times}$. Show that if $B \otimes_{\mathbb{Q}} \ell$ is split (so $(B^{\operatorname{op}} \otimes_{\mathbb{Q}} \mathbb{Q}_{\ell})^{\times} \cong \operatorname{GL}_2(\mathbb{Q}_{\ell})$), and v is a place of good reduction for A, then $\operatorname{tr} \rho_{A,\ell}(\operatorname{Frob}_v) \in \mathbb{Q}$. Why do examples of this type not occur over when $K = \mathbb{Q}$?
- 2. (*) Let K be a number field. The group SL_2 satisfies the strong approximation property: the group $\operatorname{SL}_2(K)$ is dense in $\operatorname{SL}(\mathbb{A}_K^{\infty})$ (embedded diagonally). Use this to show that for any open compact subgroup $U \subset \operatorname{GL}_2(\mathbb{A}_K^{\infty})$, the map det : $\operatorname{GL}_2 \to \operatorname{GL}_1$ induces a bijection $\pi_0(Y_U) \cong K^{\times} \setminus \mathbb{A}_K^{\infty} / \det(U)$. In particular, this set is finite (why?).
- 3. Recall that the algebra of Hamiltonian quaternions consists of all elements a+bi+cj+dk, where $a,b,c,d\in\mathbb{R}$ are central and the elements i,j,k satisfy the relations $i^2=j^2=k^2=-1$ and $ij=-ji=k,\ jk=-kj=i$, and ki=-ik=j. We can define the hyperbolic upper half-space \mathbb{H}_3 as follows: it is the set of Hamiltonian quaternions x+jy where $x\in\mathbb{C}$ and $y\in\mathbb{R}_{>0}$.
 - (a) Show that $SL_2(\mathbb{C})$ acts on \mathbb{H}_3 by the formula

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \tau = (A\tau + B)(C\tau + D)^{-1}.$$

Show that we can extend this to an action of $GL_2(\mathbb{C})$, by making $\mathbb{C}^{\times} \subset GL_2(\mathbb{C})$ act trivially.

- (b) Show that there is an isomorphism $\operatorname{GL}_2(\mathbb{C})/\operatorname{U}(2)\mathbb{R}^\times\cong \mathbb{H}_3$ of homogeneous spaces for $\operatorname{GL}_2(\mathbb{C})$. Deduce that if K is an imaginary quadratic field of class number one, and $U=\operatorname{GL}_2(\widehat{\mathcal{O}}_K)$, then there is an isomorphism $Y_U\cong\operatorname{GL}_2(\mathcal{O}_K)\backslash\mathbb{H}_3$. This is an example of a *Bianchi manifold* (or orbifold).
- 4. Let \mathcal{J} denote the set of open compact subgroups of $GL_2(\mathbb{A}_K^{\infty})$. The group $GL_2(\mathbb{A}_K^{\infty})$ acts on \mathcal{J} by conjugation. If $U \in \mathcal{J}$, $g \in GL_2(\mathbb{A}_K^{\infty})$, then there is a map $Y_{gUg^{-1}} \to Y_U$ given on elements $(h,x) \in GL_2(\mathbb{A}_K^{\infty}) \times X$ by the formula $(h,x) \mapsto (hg,x)$. Use this to construct a structure on

$$\mathcal{A} = \varinjlim_{U} H^*(Y_U, \mathbb{C})$$

of (left) $\mathbb{C}[\operatorname{GL}_2(\mathbb{A}_K^{\infty})]$ -module.

3.3 Exercises to lecture 3

- 1. Complete the proof of Artin formalism.
- 2. Let E be an elliptic curve over \mathbb{Q} with complex multiplication. Use Artin formalism and known properties of Hecke L-functions to show that L(E,s) admits an analytic continuation and satisfies a functional equation.
- 3. (*) Brauer's theorem states that for any finite group G, the Grothendieck group of $\mathbb{C}[G]$ is generated by representations of the form $\operatorname{Ind}_H^G \chi$, where $\chi: H \to \mathbb{C}^\times$ is a character and $H \subset G$ is an elementary subgroup (i.e. of the form $H = C \times P$, where C is cyclic and P is a p-group for some prime p).

By definition, an Artin representation is a continuous representation $\rho: G_K \to GL_n(\mathbb{C})$ of finite image. Use Brauer's induction theorem to show that for any such representation, $L(\rho, s)$ admits a meromorphic continuation to the complex plane and satisfies a functional equation. What is the order of the pole at s = 1?

4. Show by example that we cannot use the technique in the previous exercise to show that $L(\rho, s)$ admits an analytic continuation to the complex plane (except for a possible pole at s=1). (Hint: visit groupnames.org and look at groups of order 24.)

3.4 Exercises to lecture 4

- 1. In this exercise, we show how to get an automorphic representation of $GL_2(\mathbb{A}_{\mathbb{Q}})$ from a modular elliptic curve E over \mathbb{Q} . Recall that we have associated to E a holomorphic differential $F(\tau)d\tau$ on \mathbb{H} which is invariant under the action of $\Gamma_0(N)$. We want to lift this to a function $\phi: GL_2(\mathbb{A}) \to \mathbb{C}$ which is invariant under left translation by $GL_2(\mathbb{Q})$ (and invariant under right translation by $U_0(N) \subset GL_2(\mathbb{A}^{\infty})$).
 - (a) If $(g^{\infty}, g_{\infty}) \in U_0(N) \times \operatorname{GL}_2(\mathbb{R})^+ \subset \operatorname{GL}_2(\mathbb{A})$, we define $\phi(g) = \det(g_{\infty})^{3/2} f(g_{\infty}i) j(g_{\infty}, i)^{-2}$, where $j(g, \tau) = c\tau + d$ for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Show that ϕ extends uniquely to a function $\operatorname{GL}_2(\mathbb{A}) \to \mathbb{C}$ which is invariant under left translation by $\operatorname{GL}_2(\mathbb{Q})$.
 - (b) The centre of GL_2 is $Z = \mathbb{G}_m$, embedded diagonally. Show that ϕ transforms under the action of $Z(\mathbb{A}_{\mathbb{Q}})$ according to the norm character $\omega = \|\cdot\| : \mathbb{A}_{\mathbb{Q}}^{\times} \to \mathbb{R}_{>0}$.
 - (c) Show that ϕ is cuspidal. (Hint: try to interpret this in terms of a property of the function F.)

The automorphic representation corresponding to E is the subrepresentation of $L^2_{\omega,0}(\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}))$ generated by ϕ .

2. (*) Classify the conjugacy classes of continuous semisimple representations $\phi: W_{\mathbb{R}} \to \mathrm{GL}_2(\mathbb{C})$. Which ones have the property that $\phi|_{\mathbb{C}^{\times}}$ is algebraic?