Modularity of Elliptic Curves Defined over the Rationals

Chloe Martindale

April 13, 2016

Contents

1 What is modularity? 1

2 Representations and Modularity 2

3 The main theorems required to prove modularity of elliptic curves over $\mathbb{Q}$ 6

These are notes from a talk presented in the Elliptic Curves Seminar at Universiteit Leiden in April 2016, organised by Peter Bruin and Manolis Tzortzakis. The purpose of these notes is to explain (to some extent) the beautiful proof of Fermat's Last Theorem given by Andrew Wiles in [Wil95], by means of showing that every semistable elliptic curve over $\mathbb{Q}$ is modular, and the generalisation of Wiles' method by Breuil, Conrad, Diamond and Taylor in [BCDT01] proving that *every* elliptic curve over $\mathbb{Q}$ is modular. There has been a lot of work in this area to generalise Wiles' method to prove modularity of elliptic curves over more general fields, but unfortunately we will not have time to cover that here.

1 What is modularity?

We first recall the definition of a modular curve of level $N \in \mathbb{Z}_{>0}$, as it is needed to define modularity of an elliptic curve.

Definition 1.1. For $N \in \mathbb{Z}_{>0}$, let

$$\Gamma_0(N) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod N \right\}.$$

Then $\Gamma_0(N)$ acts on $\mathbb{H}$, the complex upper half plane, by Möbius transformations, so that we can define

$$Y_0(N) := \Gamma_0(N) \backslash \mathbb{H}.$$

This is non-compact, but it can be compactified by adding a finite number of points (called cusps), given by

$$\Gamma_0(N)\backslash\mathbb{P}^1(\mathbb{Q}).$$

For this talk, we will define the compact projective algebraic curve

$$X_0(N) := Y_0(N) \cup \Gamma_0(N)\backslash\mathbb{P}^1(\mathbb{Q})$$

to be the *modular curve of level* N.

Definition 1.2. If an elliptic curve E over $\mathbb{Q}$ has a finite covering by a modular curve $X_0(N)$ then we say that E is *modular*.

There is a rich history behind the study of modular elliptic curves, some of which we now recall:

- In the 1950s and 60s, Shimura and Taniyama conjectured that every elliptic curve is modular.
- In 1985, Frey observed that if the Shimura-Taniyama conjecture holds for semistable elliptic curves over $\mathbb{Q}$, this may prove Fermat's Last Theorem, and in 1986 Ribet proved that this is the case.
- In 1995, Wiles proved in [Wil95] that every semistable elliptic curve over $\mathbb{Q}$ is modular, thus proving Fermat's Last Theorem. He also observed that the techniques he used could be generalised.
- In 1999, Conrad, Diamond and Taylor proved in [CDT99] that every elliptic curve over $\mathbb{Q}$ such that 27 does not divide its conductor is modular, generalising the techniques developed by Wiles.
- In 2001, Breuil, Conrad, Diamond and Taylor proved in [BCDT01] that every elliptic curve over $\mathbb{Q}$ is modular, extending their previous work.

One important consequence of an elliptic curve being modular is that its L-function $L(E, s)$ extends to a meromorphic function for all complex s, which will be important for Raymond van Bommel's talk on the Birch and Swinnerton-Dyer Conjecture next week. We now recall some representation theory and define what it means for a representation to be modular (this turns out to be key in proving that elliptic curves are modular by considering representations associated to them).

2 Representations and Modularity

We first recall some representation theory as it is essential for later sections.

Definition 2.1.

1. Let G be a group and k a field. A *representation* of G over k is a group homomorphism
$$\rho : G \longrightarrow \mathrm{GL}(V)$$
for some k-vector space V.

2. Let k be a field and A a k-algebra. A *representation* of A is a k-algebra homomorphism
$$\rho : A \longrightarrow \mathrm{End}_k(V)$$
for some k-vector space V.

We can associate representations to elliptic curves; we now define the ℓ-adic Galois representation of an elliptic curve.

Definition 2.2. The *ℓ-adic Tate module* of an elliptic curve E over $\mathbb{Q}$ is given by
$$T_\ell(E) = \lim_{\infty \leftarrow n} E[\ell^n],$$
where the inverse limit is taken with respect to multiplication by ℓ.

The ℓ-adic Tate module is a free $\mathbb{Z}_\ell$-module of rank 2. Now $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts continuously on $E[\ell^n]$, and this action commutes with multiplication by ℓ, so that in particular there exists a representation
$$\rho_{E,\ell} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{Aut}(T_\ell(E)).$$
This is defined to be the *ℓ-adic Galois representation of E*. We will define the notion of modularity of Galois representations, since via the above representation this will eventually give us modularity of elliptic curves. We first need to define some other properties of representations.

Definition 2.3. Let G be a group, let k be a field and let $\rho : G \to \mathrm{GL}(V)$ be a representation of G over k. Then we say that ρ is *reducible* if there is a propoer non-trivial k-vector subspace W of V such that for every $g \in G$,
$$\rho(g)W \subseteq W.$$
Otherwise, we say that ρ is *irreducible*. If every lift of ρ to a representation of G over $\overline{k}$, the algebraic closure of k, is irreducible, we say that ρ is *absolutely irreducible*.

Definition 2.4. Let G be a group and k a field. For k-vector spaces V_1 and V_2 and representations $\rho_i : G \to \mathrm{GL}(V_i)$ for $i = 1, 2$, we say that ρ_1 and ρ_2 are *equivalent* if there exists a k-vector space isomorphism
$$\alpha : V_1 \longrightarrow V_2$$
such that for every $g \in G$,
$$\alpha \circ \rho(g) \circ \alpha^{-1} = \rho'(g).$$

Definition 2.5. Let K be a local field, and $K \subset L$ a Galois extension of K. The *inertia group* $I_{L/K}$ is defined to be

$$I_{L/K} = \{\sigma \in \mathrm{Gal}(L:K) : \forall x \in \mathcal{O}_L,\, \sigma(x) \equiv x \bmod \mathfrak{m}_L\},$$

where $\mathcal{O}_L$ denotes the valuation ring of L (i.e. all the elements of L that have non-negative valuation) and $\mathfrak{m}_L$ is the maximal ideal of $\mathcal{O}_L$.

Exercises 2.6.

1. Suppose that $K = \mathbb{Q}_p$, where p is a rational prime that is inert in $\mathbb{Q}_p(\sqrt{2})$, and suppose that $L = \mathbb{Q}_p(\sqrt{2})$. Show that $I_{L/K} = \mathrm{Id}$.
2. Show that $I_{\mathbb{Q}_p(\sqrt{p})/\mathbb{Q}_p} = \mathrm{Gal}(\mathbb{Q}_p(\sqrt{p})/\mathbb{Q}_p) = C_2$.

Definition 2.7. Let ℓ be a rational prime, let V be a $\mathbb{Q}_\ell$-vector space and let $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(V)$ be a representation. Let $p \neq \ell$ be a rational prime. For $\mathfrak{p}$ a prime in $\overline{\mathbb{Q}}$ above p, we define $I_\mathfrak{p}$ to be the image of $I_{\overline{\mathbb{Q}}_p/\mathbb{Q}_p}$ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ under the natural embedding

$$\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \hookrightarrow \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}).$$

We say that ρ is *unramified* at p if for every $\mathfrak{p}$ above p, we have

$$I_\mathfrak{p} \in \ker(\rho).$$

We have now defined what it means for a representation to be reducible, irreducible, absolutely irreducible and unramified, and when 2 representations are equivalent. In order to define what it means for a representation to be modular, we must first define cusp forms. To this end, replacing $\Gamma_0(N)$ in Theorem 1.1 by

$$\Gamma_1(N) := \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod N\right\},$$

we define the compact projective algebraic curve $X_1(N)$ in the same way that we defined $X_0(N)$.

Definition 2.8. A *cusp form* of weight $k \in \mathbb{Z}_{\geq 1}$ and level $N \in \mathbb{Z}_{\geq 1}$ is a holomorphic function

$$f : \mathbb{H} \longrightarrow \mathbb{C}$$

such that

1. for every $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(N)$ and for every $z \in \mathbb{H}$, we have

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z),$$

and

2. $|f(z)|^2(\mathrm{Im}(z))^k$ is bounded on $\mathbb{H}$.

we denote the finite dimensional $\mathbb{C}$-vector space of cusp forms of weight k and level N by $S_k(N)$.

Every $f \in S_k(N)$ has a Fourier expansion given by

$$f(z) = \sum_{n=1}^{\infty} c_n(f) e^{2\pi i n z}.$$

Definition 2.9. The L-series of $f \in S_k(N)$ at $s \in \mathbb{C}$ is defined by

$$L(f, s) = \sum_{n=1}^{\infty} c_n(f)/n^s.$$

We will need the following aspects of cusp forms:

1. For every prime p that does not divide N, there exists a linear operator

   $$T_p : S_k(N) \to S_k(N)$$

   called a *Hecke operator*. (There is an explicit definition but we won't need it).

2. Operators T_p for p not dividing N can be simultaneously diagonlised on $S_k(N)$ and a simultaneous eigenvector is called an *eigenform*.

3. For f an eigenform with eigenvalues $a_p(f)$, the eigenvalues are algebraic integers and
   $$c_p(f) = a_p(f) c_1(f).$$

Now work of Deligne, Riber, Serre and Shimura gives us the following theorem, which describes how to associate a representation to a cusp form. Here all notation is as above.

Theorem 2.10. *Let λ be a place of the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$ above ℓ, and let $\overline{\mathbb{Q}}_\lambda$ be the algebraic closure of $\mathbb{Q}_\ell$ as a $\overline{\mathbb{Q}}$-algebra via λ. If $f \in S_k(N)$ is an eigenform, the there exists and unique continuous irreducible representation*

$$\rho_{f,\lambda} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\overline{\mathbb{Q}}_\lambda)$$

such that for every p not dividing $N\ell$, we have that $\rho_{f,\lambda}$ is unramified at p and

$$\mathrm{tr}(\rho_{f,\lambda}(\mathrm{Frob}_p)) = a_p(f).$$

Remark 2.11. Let $\iota : \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \hookrightarrow \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be the natural embedding. Then we define Frob_p to be the image under ι of a lift of a topological generator of $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$. For every choice of $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, the image under $\rho_{f,\lambda}$ will be the same because $\rho_{f,\lambda}$ is unramified.

Now we can define what it means for a representation to be modular!

Definition 2.12. Suppose that

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\overline{\mathbb{Q}}_\lambda)$$

is a continuous representation. If there exist $k, N \in \mathbb{Z}_{>0}$ such that there exists an eigenform $f \in S_k(N)$ and a place $\lambda|\ell$ such that

$$\rho \sim \rho_{f,\lambda},$$

then we say that ρ is *modular*. Furthermore, if a continuous representation $\overline{\rho}$ for $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ over $\mathbb{Q}_\ell$, $\overline{\mathbb{F}}_\ell$ or $\mathbb{F}_\ell$ lifts to a modular representation, then we say that $\overline{\rho}$ is *modular*.

3 The main theorems required to prove modularity of elliptic curves over $\mathbb{Q}$

Langlands and Tunnel had already proved that an irreducible representation

$$\overline{\rho}_{E,3} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{Aut}(E[3])$$

is modular, giving Wiles part of the proof. The first theorem that Wiles proves (using the Langlands-Tunnel theorem) in [Wil95] is as follows.

Theorem 3.1. *Let E be an elliptic curve over $\mathbb{Q}$. Suppose that*

1. *E has good or multiplicative reduction at 3,*
2. *$\overline{\rho}_{E,3}|_{\mathbb{Q}(\sqrt{-3})}$ absolutely irreducible, and*
3. *define $D_q = \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}_q/\mathbb{Q}_q) : \sigma(q) = q\} \leq I_q$; for any $q \equiv -1 \bmod 3$ either $\overline{\rho}_{E,3}|_{D_q}$ is reducible over the algebraic closure or $\overline{\rho}_{E,3}|_{I_q}$ is absolutely irreducible. Here I_q is the inertia group for $\overline{\mathbb{Q}}q/\mathbb{Q}_q$. Then E is modular.*

Remark 3.2. Recall that an elliptic curve over $\mathbb{Q}$ has multiplicative reduction at a prime p if the reduction of the elliptic curve mod p is singular, and its unique singular point is nodal. Recall also that an elliptic curve over $\mathbb{Q}$ is semistable if it has good or multiplicative reduction at every prime p.

Wiles uses the above theorem to deduce his main theorem below.

Theorem 3.3. *Let E be a semistable elliptic curve over $\mathbb{Q}$. Then E is modular.*

To prove this, Wiles splits into 2 cases. For the case in which $\overline{\rho}_{E,3}$ is irreducible, he just needs to show that the conditions of Theorem 3.1 are satisfied (note that this is the Langlands-Tunnel case). However, for the case in which $\overline{\rho}_{E,3}$ is reducible, Theorem 3.1 does not apply. Wiles instead shows that

$$\overline{\rho}_{E,3} \text{ reducible } \Rightarrow \overline{\rho}_{E,5} \text{ irredcuible.}$$

For the 5-adic case there is no Langlands-Tunnel, so an analogy of Theorem 3.1 is more difficult. Wiles shows instead that if $\overline{\rho}_{E,5}$ is irreducible, then there is no quadratic twist E' of E over $\mathbb{Q}(\sqrt{5})$ such that

1. $\overline{\rho}_{E,5}$ is an induced representation over $\mathbb{Q}(\sqrt{5})$, and

2. E' is semistable at 5.

He proves an analogous statement to Theorem 3.1 with these hypotheses, that is, he proves that

$$\overline{\rho}_{E,5} \text{ irreducible } \Rightarrow E \text{ modular},$$

thus proving that all semistable elliptic curves over $\mathbb{Q}$ are modular.

Breuil, Conrad, Diamond and Taylor were able to remove the condition that E is semistable by further studying the 5-adic representation E. In 1999, Conrad, Diamond and Taylor proved the following theorem.

Theorem 3.4. *Let E be an elliptic curve over $\mathbb{Q}$. If $\overline{\rho}_{E,5}$ is modular or $\overline{\rho}_{E,5}|_{\mathbb{Q}(\sqrt{5})}$ is not absolutely irreducible, then E is modular.*

From this they could deduce that every elliptic curve over $\mathbb{Q}$ with conductor not divisible by 27 is modular. The final step in showing that all elliptic curves over $\mathbb{Q}$ are modular was the theorem of Breuil, Conrad, Diamond and Taylor in 2001:

Theorem 3.5. *Any continuous absolutely irreducible representation*

$$\overline{\rho} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{F}_5)$$

with cyclotomic determinant is modular.

References

[BCDT01] C. Breuil, B. Conrad, F. Diamond, R. Taylor, *On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercsies*, J. Amer. Math. Soc. 14, no. 4 (2001), 843-939,

[CDT99] B. Conrad, F. Diamond, R. Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations*, J. Amer. Math. Soc. 12 (1999), 521-567,

[Wil95] A. J. Wiles, *Modular elliptic curves and Fermats Last Theorem*, Ann. of Math. 141 (1975), 443-551.