Elliptic Curves and Modularity

Jack A. Thorne

Abstract

We survey results and conjectures concerning the modularity of elliptic curves over number fields.

Mathematics Subject Classification 2020. Primary 11G05; Secondary 11R39

Keywords. Elliptic curves, automorphic forms, Langlands Program

Contents

1 Introduction

The modularity conjecture for elliptic curves over $\mathbf{Q}$ was stated with increasing degrees of precision by Taniyama, Shimura, and Weil in the 1950's and 60's. It admits several equivalent formulations, which are discussed in the textbook [17]. The most common asserts that given any elliptic curve E over $\mathbf{Q}$, we can find a newform $f \in S_2(\Gamma_0(N))$ with the property that for all but finitely many prime numbers p, the p^{th} Fourier coefficient $a_p(f)$ in the q-expansion $f(q) = q + \sum_{n \geq 2} a_n(f) q^n$ equals the number

$$a_p(E) = p + 1 - |E(\mathbf{F}_p)|,$$

which measures the error in the Hasse estimate for the number of points on E modulo p. The newform f is then uniquely determined by E, by the strong multiplicity one

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, UK; email: thorne@dpmms.cam.ac.uk

theorem for modular forms. Any curve E for which such a newform f exists is said to be modular.

A famous example of a modular elliptic curve is the curve of conductor 11 given by the equation

$$E : y^2 + y = x^3 - x^2.$$

This elliptic curve is modular, with associated newform

$$f(q) = q \prod_{n=1}^{\infty} (1 - q^n)^2 (1 - q^{11n})^2 \in S_2(\Gamma_0(11)).$$

The modularity conjecture is, on the face of it, a very surprising statement. It is easy to write down elliptic curves over $\mathbf{Q}$; indeed, for any cubic polynomial

$$f(x) = x^3 + ax + b \in \mathbf{Z}[x]$$

of non-zero discriminant, the equation $y^2 = f(x)$ gives an elliptic curve. On the other hand, modular forms begin life as complex analytic objects. Even once admits their algebro-geometric description (as sections of a line bundle on a modular curve, thought of as an algebraic curve over $\mathbf{Q}$), together with the theory of Hecke operators, there is no a priori reason to expect that *every* elliptic curve over $\mathbf{Q}$ should be associated to a newform. Nevertheless, the modularity conjecture was proved for semistable elliptic curves over $\mathbf{Q}$ in 1995 by Wiles and Taylor [46, 41], on the way to proving Fermat's Last Theorem, and finally for all elliptic curves over $\mathbf{Q}$ in 2001 by Breuil, Conrad, Diamond, and Taylor [8].

The modularity conjecture for elliptic curves over $\mathbf{Q}$ can be thought of as a special case of the Langlands program, in a form made precise by Clozel [13]. Newforms give rise to automorphic representations of the adèle group $\mathrm{GL}_2(\mathbf{A}_\mathbf{Q})$. Under Clozel's conjectures, there would be a correspondence between motives of rank n over a number field K (or more concretely, compatible systems of semisimple, n-dimensional representations of the absolute Galois group of K) and automorphic representations of the group $\mathrm{GL}_n(\mathbf{A}_K)$ satisfying a condition that he calls 'algebraic'. Specialising to elliptic curves, we obtain a precise analogue of the modularity conjecture valid over an arbitrary number field. (We note that such an analogue had already been anticipated, especially in the case of imaginary quadratic fields, cf. [24, 14].)

Our first goal in this article is to state the a version of this modularity conjecture for elliptic curves over a general number field K in as down to earth a manner as possible. In particular, our formulation does not use the language of automorphic representations. (This is not original; for example, Taylor's 1994 ICM article [40] contains essentially the same statement that we give here.) Note however that it is not possible to avoid the automorphic theory if one wants to give the most precise statements, or to get to the most important consequences of modularity, such as the analytic continuation of the L-function of an elliptic curve.

We will then continue to discuss some of the many applications of modularity in number theory, beyond the most famous application to Fermat's Last Theorem. It is

interesting to note that these range from statements of great theoretical importance (such as the analytic properties of the L-function) to very concrete statements that have no obvious connection to automorphic representations or the Langlands program (such as bounds on the height of solutions to Mordell's equation).

Finally, we will discuss what is known towards the modularity conjecture for elliptic curves over a number field K, beyond the case $K = \mathbf{Q}$. It is natural to break up the discussion depending on whether or not K is totally real (in the sense that each field embedding $K \to \mathbf{C}$ in fact takes values in $\mathbf{R}$). Many of the methods developed to study modularity over $\mathbf{Q}$ translate well to the totally real setting. It is more challenging to study modularity over number fields which are not totally real, but there has been much progress in this direction recently, inspired particularly by applications of Scholze's theory of perfectoid spaces.

2 The modularity conjecture

Let K be a number field, with ring of integers $\mathcal{O}_K$ and absolute Galois group $G_K = \mathrm{Gal}(\overline{K}/K)$ with respect to a fixed choice of algebraic closure $\overline{K}/K$. (More generally, if k is a perfect field then we will write G_k for the absolute Galois group of k with respect to some fixed choice of algebraic closure.)

Definition 2.1. An elliptic curve over K is a pair (E, ∞), where E is a smooth, projective, connected curve over K and $\infty \in E(K)$ is a marked rational point.

We often take the marked point as given and just say that E is an elliptic curve. Any elliptic curve may be given by a Weierstrass equation

$$y^2 = x^3 + ax + b, \tag{2.1}$$

where $a, b \in \mathcal{O}_K$ and x, y are plane co-ordinates. The closure (in the projective plane $\mathbf{P}^2$) of the affine curve defined by such an equation picks up exactly one extra point at infinity, which is the marked point ∞. The discriminant $\Delta = \Delta(a, b) = -16(4a^3 + 27b^2)$ is non-zero. Conversely, for any pair $(a, b) \in \mathcal{O}_K^2$ such that $\Delta(a, b) \neq 0$, the equation (2.1) defines an elliptic curve.

Elliptic curves have a number of important associated structures. The first is the group law: there is a unique way to make any elliptic curve into a commutative algebraic group with identity element $\infty \in E(K)$. The addition law then has a simple characterization: three points P, Q, R sum to ∞ if and only if they are collinear in the Weierstrass embedding (2.1).

The second is the system of reductions modulo v, for v a finite (i.e. non-archimedean) place of the number field K. If the discriminant Δ of a given Weierstrass equation is a v-adic unit, then v is a place of good reduction for the curve E: the reduction modulo v of the Weierstrass equation (2.1) defines an elliptic curve over the residue field $k(v)$ of the completion K_v of K at the place v. This leads to the definition of the quantity

$$a_v(E) = q_v + 1 - |E(k(v))|,$$

where $q_v = |k(v)|$ and $|E(k(v))|$ is the number of points of this reduced curve over the residue field $k(v)$. One can also define a_v at the places where Δ is not a v-adic unit, but this requires the use of a long Weierstrass equation in order to be able to find a model of minimal discriminant at the place v (see [38, Ch. VII]).

The third structure we want to introduce is the compatible system of ℓ-adic Galois representations of E. For each prime number ℓ, the étale cohomology group $H^1_{\text{ét}}(E_{\overline{K}}, \mathbf{Q}_\ell)$ is a 2-dimensional $\mathbf{Q}_\ell$-vector space which receives a continuous action of the absolute Galois group G_K. Fixing a choice of basis, we obtain a continuous representation

$$\rho_{E,\ell} : G_K \to \mathrm{GL}_2(\mathbf{Q}_\ell).$$

(This representation is the same, up to passing to the dual and taking a twist by the cyclotomic character, of the representation afforded by the ℓ-adic Tate module of E.) If v is a finite place of K not dividing ℓ and at which E has good reduction, then the representation $\rho_{E,\ell}$ is unramified at v. By definition, this means that the inertia subgroup I_{K_v} of the decomposition group $G_{K_v} \subset G_K$ acts trivially through $\rho_{E,\ell}$. Moreover, if $\mathrm{Frob}_v \in G_{K_v}/I_{K_v} \cong G_{k(v)}$ denotes the Frobenius element[2] then we have the equality

$$\mathrm{tr}\rho_{E,\ell}(\mathrm{Frob}_v) = a_v(E),$$

a consequence of the Grothendieck–Lefschetz trace formula for the reduction modulo v of the elliptic curve E. We call the collection $(\rho_{E,\ell})_\ell$ of ℓ-adic representations a 'compatible system' because these Frobenius traces are independent of ℓ (even though the representations themselves are incomparable, because the topological fields $\mathbf{Q}_\ell$ are pairwise non-isomorphic).

So much for elliptic curves. We next want to introduce the structures 'on the automorphic side' that should be matched up with elliptic curves under the modularity conjecture. By analogy with class field theory, which gives a description of the 1-dimensional representations of G_K, these structures should be defined using the 'internal arithmetic' of the field K. To write these down, we first need to recall the existence of the adèle ring of K.

Definition 2.2. The finite adèle ring of K is the restricted direct product

$$\mathbf{A}_K^\infty = \prod_{v \text{ finite}}{}' K_v$$

with respect to the valuation rings $\mathcal{O}_{K_v} \subset K_v$. The adèle ring of K is the product $\mathbf{A}_K = \mathbf{A}_K^\infty \times K_\infty$, where $K_\infty = \prod_{v \text{ infinite}} K_v$.

In other words, $\mathbf{A}_K$ is the set of elements $x = (x_v)_v \in \prod_v K_v$ such that for all but finitely many finite places v of K, $x_v \in \mathcal{O}_{K_v}$. The fundamental facts concerning $\mathbf{A}_K$ are: it is a locally compact topological ring, the diagonal embedding $K \to \mathbf{A}_K$ induces the discrete topology on K, and the quotient $\mathbf{A}_K/K$ is compact.

[2] More precisely, the geometric Frobenius element (inverse of the arithmetic Frobenius automorphism $x \mapsto x^{q_v}$ on $\overline{k(v)}$).

Having defined $\mathbf{A}_K$, we can take the $\mathbf{A}_K$-points of any algebraic group over K. In particular, the group $\mathrm{GL}_2(\mathbf{A}_K)$ is then defined. This group can also be realised as the restricted direct product $\prod'_v \mathrm{GL}_2(K_v)$, with respect to the family of open subgroups $\mathrm{GL}_2(O_{K_v}) \subset \mathrm{GL}_2(K_v)$ for finite places v.

Definition 2.3. Let $\mathfrak{n} \subset O_K$ be a non-zero ideal. We define the open compact subgroup of $\prod_{v \text{ finite}} \mathrm{GL}_2(O_{K_v})$

$$U_1(\mathfrak{n}) = \left\{ \begin{pmatrix} a_v & b_v \\ c_v & d_v \end{pmatrix} \in \prod_{v \text{ finite}} \mathrm{GL}_2(O_{K_v}) : \forall v, c_v, d_v - 1 \equiv 0 \bmod \mathfrak{n} O_{K_v} \right\}.$$

If v is an infinite place of K, we let $U_v = \mathrm{SO}_2(\mathbf{R})$ (if $K_v = \mathbf{R}$) or $U_v = \mathrm{U}_2(\mathbf{R})$ (if $K_v \cong \mathbf{C}$). Let $U_\infty = \mathbf{R}_{>0} \cdot \prod_{v|\infty} U_v \subset \mathrm{GL}_2(K_\infty)$. We then define the quotient topological space

$$Y_1(\mathfrak{n}) = \mathrm{GL}_2(K) \backslash \mathrm{GL}_2(\mathbf{A}_K) / U_1(\mathfrak{n}) \times U_\infty.$$

In order to formulate the modularity conjecture, we will look at the singular cohomology groups $H^*(Y_1(\mathfrak{n}), \mathbf{Q})$. These are finite dimensional $\mathbf{Q}$-vector spaces. Indeed, $Y_1(\mathfrak{n})$ can be represented quite concretely, as we now explain. The double quotient $\mathrm{GL}_2(K) \backslash \mathrm{GL}_2(\mathbf{A}_K^\infty) / U_1(\mathfrak{n})$ (where we omit the infinite places) is finite; if $g_1, \ldots, g_n \in \mathrm{GL}_2(\mathbf{A}_K^\infty)$ are coset representatives, then $Y_1(\mathfrak{n})$ can itself be identified with the disjoint union of the quotients $\Gamma_i \backslash \mathrm{GL}_2(K_\infty) / U_\infty$, where we define

$$\Gamma_i = \mathrm{GL}_2(K) \cap g_i U_1(\mathfrak{n}) g_i^{-1}$$

(intersection in $\mathrm{GL}_2(\mathbf{A}_K^\infty)$). The groups Γ_i are congruence subgroups of $\mathrm{GL}_2(K)$, which are torsion-free if the ideal $\mathfrak{n}$ is small enough, so these quotients are generalisations of the modular curves arising in the theory of classical modular forms. In fact, if $K = \mathbf{Q}$ and $\mathfrak{n} = (N)$ for a natural number N, then the space $Y_1(\mathfrak{n})$ defined above may be identified with the usual modular curve of level $\Gamma_1(N)$.

The reason for defining $Y_1(\mathfrak{n})$ using the adèle ring is that it makes transparent the definition of Hecke operators, which are necessary to be able to give a precise formulation of the modularity conjecture. The existence of Hecke operators is a consequence of the following observation: if $U \subset \mathrm{GL}_2(\mathbf{A}_K^\infty)$ is any open compact subgroup, let Y_U be the space defined in the same way as $Y_1(\mathfrak{n})$, except with $U_1(\mathfrak{n})$ replaced by U. If $V \subset U$ then there is a natural projection $Y_V \to Y_U$. We can thus form the direct limit

$$\mathcal{A} = \varinjlim_U H^*(Y_U, \mathbf{Q}),$$

a representation of $\mathrm{GL}_2(\mathbf{A}_K^\infty)$ which is *smooth*, in the sense that each vector is fixed by some open compact subgroup of $\mathrm{GL}_2(\mathbf{A}_K^\infty)$. Moreover, we can recover $H^*(Y_1(\mathfrak{n}), \mathbf{Q})$ as the space of $U_1(\mathfrak{n})$-invariant vectors of $\mathcal{A}$. General considerations (see e.g. [33, §2.2]) then imply that $H^*(Y_1(\mathfrak{n}), \mathbf{Q})$ has the structure of module for the ring $\mathcal{H}(\mathrm{GL}_2(\mathbf{A}_K^\infty), U_1(\mathfrak{n}))$ of compactly supported, $U_1(\mathfrak{n})$-biinvariant functions $f : \mathrm{GL}_2(\mathbf{A}_K^\infty) \to \mathbf{Q}$. Elements of this ring are what we call Hecke operators.

The most fundamental ones are as follows:

Definition 2.4. Let v be a finite place of K which is prime to $\mathfrak{n}$, and let $\varpi_v \in \mathcal{O}_{K_v}$ be a uniformizer of the valuation ring at v. Then the Hecke operator

$$T_v : H^*(Y_1(\mathfrak{n}), \mathbf{Q}) \to H^*(Y_1(\mathfrak{n}), \mathbf{Q})$$

is the endomorphism induced by the characteristic function $f_v \in \mathcal{H}(\mathrm{GL}_2(\mathbf{A}_K^\infty), U_1(\mathfrak{n}))$ of the double coset $U_1(\mathfrak{n})xU_1(\mathfrak{n})$, where $x = (x_w)_w \in \mathrm{GL}_2(\mathbf{A}_K^\infty)$ is the element with $x_w = 1$ if $w \neq v$ and $x_w = \mathrm{diag}(\varpi_v, 1)$ if $w = v$.

This definition is independent of the choice of uniformizer x_v. The Hecke operator T_v can also be described more concretely as the endomorphism of $H^*(Y_1(\mathfrak{n}), \mathbf{Q})$ induced by a correspondence

$$\begin{array}{ccc} & Y_{U_1(\mathfrak{n}) \cap xU_1(\mathfrak{n})x^{-1}} & \\ \swarrow & & \searrow \\ Y_1(\mathfrak{n}) & & Y_1(\mathfrak{n}) \end{array}$$

However, its definition is explained most clearly by the local Langlands correspondence for unramified representations of $\mathrm{GL}_2(K_v)$, as we will recall below.

We now have everything we need to state a version of the modularity conjecture:

Conjecture 2.5. *Let E be an elliptic curve over K such that* $\mathrm{End}_K(E) = \mathbf{Z}$. *Then there exists an ideal $\mathfrak{n} \subset \mathcal{O}_K$ and a non-zero class $c_E \in H^*(Y_1(\mathfrak{n}), \mathbf{Q})$ such that for all but finitely many finite places v of K, we have the equality*

$$T_v(c_E) = a_v(E)c_E.$$

Various remarks are in order. The restriction to curves with $\mathrm{End}_K(E) = \mathbf{Z}$ is made because curves with $\mathrm{End}_K(E) \neq \mathbf{Z}$ (in other words, elliptic curves with complex multiplication defined over K) behave differently: their Galois representations $\rho_{E,\ell}$ are abelian, and are described by class field theory. We note that the condition $\mathrm{End}_K(E) = \mathbf{Z}$ always holds if K is totally real, for example if $K = \mathbf{Q}$.

Next we ask how this conjecture is related to the more classical conjecture in the case $K = \mathbf{Q}$ referenced in the introduction, which phrases modularity in terms of the modular forms, rather than cohomology classes. The bridge between modular forms and cohomology is in this case given by the Eichler–Shimura isomorphism. This is an isomorphism

$$M_2(\Gamma_1(N)) \oplus \overline{S_2(\Gamma_1(N))} \cong H^1(Y_1(N), \mathbf{C})$$

respecting the action of Hecke operators on each side. If p is a prime number not dividing N and f is a newform, then the p^{th} Fourier coefficient of f coincides with the eigenvalue of the Hecke operator T_p on f, which explains how newforms give rise to cohomology classes in $H^1(Y_1(N), \mathbf{C})$ which are eigenvectors for Hecke operators. When the eigenvalues are rational numbers, we can even choose eigenvectors which lie in $H^1(Y_1(N), \mathbf{Q})$.

How is this conjecture related to the formulation of Clozel [13], also referenced in the introduction, which would lead one to associate to each elliptic curve E over K with $\mathrm{End}_K(E) = \mathbf{Z}$ a cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbf{A}_K)$? Such a representation π admits a restricted tensor product decomposition $\pi = \otimes'_v \pi_v$, indexed by the set of places v of the number field K. One can predict the isomorphism class of π_v, as a representation of the group $\mathrm{GL}_2(K_v)$, using the local Langlands correspondence for the group $\mathrm{GL}_2(K_v)$. Let us recall that when v is a finite place, the local Langlands correspondence rec_{K_v} is a bijection between the following two sets of objects:

- The set of isomorphism classes of irreducible smooth representations of $\mathrm{GL}_2(K_v)$ over $\mathbf{C}$.
- The set of isomorphism classes of 2-dimensional Frobenius-semisimple Weil–Deligne representations of the Weil group $W_{K_v} \subset G_{K_v}$ over $\mathbf{C}$.

We can use the local Langlands correspondence to build an irreducible representation $\pi(E)$ of $\mathrm{GL}_2(\mathbf{A}_K^\infty)$ from an elliptic curve E over K, by specifying a Weil–Deligne representation (r_v, N_v) of the group W_{K_v} for each finite place v of K using the local representations $\rho_{E,\ell}|_{W_{K_v}}$. (For an explanation of how to do this, see e.g. [39]). Thus $\pi(E)$ is the restricted tensor product of the local factors $\mathrm{rec}_{K_v}^{-1}(r_v \otimes |\cdot|^{1/2}, N_v)$. In particular, this leads to the following more precise conjecture, which implies Conjecture 2.5:

Conjecture 2.6. *Let E be an elliptic curve over K such that $\mathrm{End}_K(E) = \mathbf{Z}$, and let $\pi(E)$ be the irreducible smooth representation of $\mathrm{GL}_2(\mathbf{A}_K^\infty)$ associated to E using the local Langlands correspondence. Then there is a $\mathrm{GL}_2(\mathbf{A}_K^\infty)$-equivariant injection $\pi(E) \hookrightarrow \mathcal{A} \otimes_{\mathbf{Q}} \mathbf{C}$.*

From this point of view we can explain the importance of the Hecke operators T_v in formulating the modularity conjecture, which is otherwise slightly obscure: if v is a finite place of K, then the local Langlands correspondence restricts to a bijection between the following two sets of objects:

- The set of isomorphism classes of smooth representations of $\mathrm{GL}_2(K_v)$ over $\mathbf{C}$ which are *unramified*, in the sense that the space of $\mathrm{GL}_2(\mathcal{O}_{K_v})$-invariant vectors is non-zero.
- The set of isomorphism classes of 2-dimensional semisimple representations of W_{K_v} which are unramified, in the sense that the inertia group $I_{K_v} \subset W_{K_v}$ acts trivially.

If π_v is an unramified irreducible smooth representation of $\mathrm{GL}_2(K_v)$ and $r \otimes |\cdot|^{1/2} = \mathrm{rec}_{K_v}(\pi_v)$, then the Hecke operator T_v acts by a scalar on the space of $\mathrm{GL}_2(\mathcal{O}_{K_v})$-invariant vectors of π_v which is equal to $\mathrm{tr}\, r(\mathrm{Frob}_v)$. We have already observed that if v is a place of good reduction for the elliptic curve E then the Grothendieck–Lefschetz

trace formula implies the equality $\rho_{E,\ell}(\mathrm{Frob}_v) = a_v(E)$, provided v is prime to ℓ. This explains the essential equality

$$\text{eigenvalue of } T_v = a_v(E)$$

which appears in the statement of Conjecture 2.5.

One can (and should) go further than we do here. For example, is it possible to describe all of the systems of Hecke eigenvalues which appear in $H^*(Y_1(\mathfrak{n}), \mathbf{C})$ in terms of abelian varieties? They cannot all be described in terms of elliptic curves since, for example, there are systems of Hecke eigenvalues which are not all rational numbers, so can not come from elliptic curves. See [40] for a precise conjectural description in terms of 'false generalised elliptic curves'.

3 Applications of modularity

We briefly discuss some applications of the modularity conjecture for elliptic curves. Our intent here is not to be exhaustive but rather to give a flavour of some of the many different applications of modularity that exist.

We mention first applications to Fermat's Last Theorem and other Fermat-style problems. Let us recall the strategy to prove Fermat's Last Theorem used in [46]. Let $p \geq 5$ be a prime number, and suppose given a non-trivial solution

$$a^p + b^p = c^p$$

to the Fermat equation in exponent p; thus $a, b, c \in \mathbf{Z}$ are coprime non-zero integers. One associates to such a non-trivial solution the Frey–Hellegoarch elliptic curve

$$E_{a,b,c} : y^2 = x(x - a^p)(x + b^p).$$

After possibly permuting a, b, c (in order to optimise the local behaviour at the prime 2), the minimal discriminant of this elliptic curve over $\mathbf{Q}$ is $2^{-8}(abc)^{2p}$ (see for example the calculation in [37, §4.1]). This implies that the reduction of the p-adic Galois representation $\rho_{E_{a,b,c},p}$ (to be discussed further below) can be ramified only at the prime 2 (and is finite flat at p). The modularity of the curve $E_{a,b,c}$, together with Ribet's level-lowering theorem, then imply the existence of a newform of weight 2 and level $\Gamma_0(2)$, a contradiction.

Variants of this strategy may be employed to study the generalised Fermat equations

$$a^p + b^q = c^r,$$

where $p, q, r \geq 2$ are integers satisfying $1/p + 1/q + 1/r < 1$. Bennett et al. [4] describe a broad range of exponents for which it can be proved using variants of the above modularity-based method that no non-trivial solutions exist. One can also study solutions to these equations in number fields other than $\mathbf{Q}$. Assuming a strengthened

version of the modularity Conjecture 2.5 for an imaginary quadratic field $K = \mathbf{Q}(\sqrt{-d})$, where $d > 0$ is an even squarefree integer, Şengün and Siksek [36] prove that for all sufficiently large prime numbers p, there are no non-trivial solutions to the Fermat equation in exponent p over O_K. See also [20] for similar (and unconditional) results over real quadratic fields.

These kinds of modular techniques can also be used to get positive (as opposed to non-existence) information about solutions to Diophantine equations. An example is given by the following theorem, taken from the work of von Känel and Matschke [26]:

Theorem 3.1. *Let a be a non-zero integer. Then any solution $(x, y) \in \mathbf{Z}^2$ to the equation $y^2 = x^3 + a$ satisfies the estimate*

$$\max(\log|x|, \frac{2}{3}\log|y|) \leq 1728|a|(\log|a| + 4).$$

Modularity is also of great importance for studying individual elliptic curves. For example, essentially all known results towards the Birch–Swinnerton-Dyer (BSD) conjecture are restricted to the class of modular elliptic curves. The BSD conjecture concerns the L-function of an elliptic curve over a number field:

Definition 3.2. Let E be an elliptic curve over a number field K. The L-function $L(E, s)$ of E is the function of a complex variable s defined by the Euler product, indexed by finite places v of K:

$$L(E, s) = \prod_{v \text{ bad}} (1 - a_v(E)q_v^{-s})^{-1} \prod_{v \text{ good}} (1 - a_v(E)q_v^{-s} + q_v^{1-2s})^{-1}.$$

The Hasse estimate implies that this Euler product converges absolutely in the right half-plane $\mathrm{Re}(s) > 3/2$, and defines a complex analytic function there. We then have the following fundamental conjectures:

Conjecture 3.3. *Let E be an elliptic curve over a number field K.*

1. *(Analytic continuation) The function $L(E, s)$ admits an analytic continuation to the whole complex plane. Defining $\Lambda(E, s) = (2\pi^{-s}\Gamma(s))^{[K:\mathbf{Q}]}L(E, s)$, there is a natural number N and a sign $\epsilon \in \{\pm 1\}$ such that the functional equation*

$$\Lambda(E, s) = \epsilon N^{1-s}\Lambda(E, 2 - s)$$

holds.

2. *(Weak BSD) Assuming (1), the order of vanishing of $L(E, s)$ at the point $s = 1$ is equal to the rank r_E of the finitely generated abelian group $E(K)$.*

3. *(Strong BSD) Assuming (2), we have*

$$\lim_{s \to 1} \frac{L(E, s)}{(s - 1)^{r_E}} = P(E)R(E)|Sha(E)|,$$

where $P(E)$ is a product of local terms, $R(E)$ is the regulator of $E(K)$, and $Sha(E)$ is the Tate–Shafarevich group of E. In particular, $Sha(E)$ is finite.

Here we follow the formulation of the strong BSD conjecture given by Gross [23], to which we refer for the definition of the terms $P(E)$, $R(E)$.

Theorem 3.4. *Let E be an elliptic curve over a number field K. Then:*

1. *If E satisfies Conjecture 2.6, then $L(E, s)$ has an analytic continuation.*
2. *If E satisfies Conjecture 2.6 and K is totally real, and either $[K : \mathbf{Q}]$ is odd or there exists a finite place v such that the Weil–Deligne representation of W_{K_v} associated to E is indecomposable, then the weak BSD conjecture holds for E provided that the order of vanishing of $L(E, s)$ at the point $s = 1$ is* at most 1.

If E satisfies Conjecture 2.6, then there is a cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbf{A}_K)$ such that $L(E, s) = L(\pi, s)$. In other words, we may identify $L(E, s)$ with an automorphic L-function. The analytic continuation of $L(E, s)$ is then a consequence of the known continuation for such automorphic L-functions [25]. When $K = \mathbf{Q}$ and $L(E, s)$ vanishes to order at most 1, the validity of the weak BSD conjecture follows from the Gross–Zagier formula and work of Kolyvagin [22, 30].

These results were generalised to a general totally real field K by Zhang [47]. It is interesting to note that the Gross–Zagier formula and its generalisations depend on the existence of a modular parameterisation, i.e. a non-constant map from a Shimura curve defined over K to the elliptic curve E. The existence of such a parameterisation for a curve E satisfying the hypothesis of Theorem 3.4 (2) is a non-trivial consequence of its modularity, in the sense of Conjecture 2.5.

4 Known results

We now discuss what is known towards the modularity Conjecture 2.5. First, it is known for elliptic curves over $\mathbf{Q}$ [46, 41, 8]:

Theorem 4.1. *Every elliptic curve over $\mathbf{Q}$ is modular.*

We review the structure of the proof, which underlies all known generalisations of this theorem. First, we change our point of view slightly by considering the modularity of the Galois representations $\rho_{E,\ell} : G_K \to \mathrm{GL}_2(\mathbf{Q}_\ell)$ associated to an elliptic curve over a number field K. For example, we can make the following definition:

Definition 4.2. Let K be a number field, let ℓ be a prime number, and let $\rho : G_K \to \mathrm{GL}_2(\mathbf{Q}_\ell)$ be a continuous representation. We say that ρ is modular if there exists a non-zero ideal $\mathfrak{n} \subset \mathcal{O}_K$ and a non-zero class $c_\rho \in H^*(Y_1(\mathfrak{n}), \mathbf{Q}_\ell)$ satisfying the following condition: for all but finitely many finite places v of K, $\rho|_{G_{K_v}}$ is unramified, c_ρ is an eigenvector of the Hecke operator T_v, and we have the equality

$$T_v(c_\rho) = (\mathrm{tr}\rho(\mathrm{Frob}_v))c_\rho.$$

In view of the equality $a_v(E) = \mathrm{tr}\rho_{E,\ell}(\mathrm{Frob}_v)$ for prime-to-ℓ places at which E has good reduction, we see that Conjecture 2.5 holds for an elliptic curve E over K if and only if one (or equivalently, all) of its ℓ-adic Galois representations is modular in the above sense.

It is important to note that this notion of modularity is very restrictive. It is believed (and known, in many cases) that any Galois representation which is modular in the above sense must be of weight 2, in the sense defined in [40]. To encompass all (say irreducible 2-dimensional) Galois representations which arise from the étale cohomology of algebraic varieties over K we would need to consider a broader definition of modularity, encompassing all of the algebraic automorphic representations of $\mathrm{GL}_2(\mathbf{A}_K)$ singled out in [13].

We can also define a notion of modularity for representations with coefficients in $\mathbf{F}_\ell$, the field with ℓ elements:

Definition 4.3. Let $\overline{\rho} : G_K \to \mathrm{GL}_2(\mathbf{F}_\ell)$ be a continuous representation. We say that $\overline{\rho}$ is modular if there exists a non-zero ideal $\mathfrak{n} \subset \mathcal{O}_K$ and a non-zero class $c_\rho \in H^*(Y_1(\mathfrak{n}), \mathbf{F}_\ell)$ satisfying the following condition: for all but finitely many finite places v of K, $\rho|_{G_{K_v}}$ is unramified, c_ρ is an eigenvector of the Hecke operator T_v, and we have the equality

$$T_v(c_\rho) = (\mathrm{tr}\rho(\mathrm{Frob}_v))c_\rho.$$

Any continuous representation $\rho : G_K \to \mathrm{GL}_2(\mathbf{Q}_\ell)$ may be conjugated to take values in $\mathrm{GL}_2(\mathbf{Z}_\ell)$; reduction modulo ℓ then gives a representation valued in $\mathrm{GL}_2(\mathbf{F}_\ell)$. We write $\overline{\rho} : G_K \to \mathrm{GL}_2(\mathbf{F}_\ell)$ for the semisimplification of this representation, which is (up to isomorphism) independent of any choices. It is easy to prove that if ρ is modular in the sense of Definition 4.2, then $\overline{\rho}$ is modular in the sense of Definition 4.3.

A fundamental idea behind the proof of Theorem 4.1 and its generalisations, first introduced in [46], is that of the modularity lifting theorem, which gives conditions under which one can go in the other direction and 'lift' the modularity of the residual representation $\overline{\rho}$ to the characteristic 0 representation ρ. Many such results now exist in the literature, all approximations to the following ideal:

Theorem Schema 4.4. *Let $\rho : G_K \to \mathrm{GL}_2(\mathbf{Q}_\ell)$ be a continuous representation satisfying the following:*

1. *Some global conditions on $\overline{\rho}$, such as the irreducibility of $\overline{\rho}$.*
2. *Some necessary local conditions on ρ, such as that ρ be of weight 2, in the sense of [40].*
3. *$\overline{\rho}$ is modular.*

Then ρ is modular.

The first such theorem, proved in [46, 41], was sufficient to establish the modularity of semistable elliptic curves over $\mathbf{Q}$ (i.e. those elliptic curves with good or multiplicative

reduction everywhere). In order to apply such a theorem, say to prove the modularity of an elliptic curve E, one needs a way to verify the modularity of the residual representation $\overline{\rho}_{E,\ell}$ for some prime ℓ. Wiles was able to do this when $\ell = 3$ and $K = \mathbf{Q}$ by exploiting a few very happy coincidences:

- The homomorphism $\mathrm{GL}_2(\mathbf{Z}_3) \to \mathrm{GL}_2(\mathbf{F}_3)$ given by reduction modulo 3 splits. Consequently, for any elliptic curve E over $\mathbf{Q}$ we can find a representation $\widetilde{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_3)$ with *finite image* and lifting $\overline{\rho}_{E,3}$.
- The group $\mathrm{GL}_2(\mathbf{F}_3)$ is soluble. The Langlands–Tunnell theorem [45], which gives the automorphy (in the sense of [13]) of 2-dimensional representations of $G_{\mathbf{Q}}$ (or more generally G_K, where K is any number field) with finite soluble image, implies that $\widetilde{\rho}$ may be associated to a weight 1 holomorphic newform.
- There exist plentiful congruences between weight 1 newforms and weight 2 newforms (for example, given by multiplying by a well-chosen weight 1 Eisenstein series). The existence of such congruences is needed to pass from the automorphy of $\widetilde{\rho}$ to the modularity of $\overline{\rho}$ in our sense (which is also the sense required for application of the modularity lifting theorem in [46]).

Verifying the modularity of $\overline{\rho}_{E,3}$ in this way, Wiles was able to prove the modularity of those semistable elliptic curves over $\mathbf{Q}$ for which $\overline{\rho}_{E,3}$ is irreducible. To take care of those curves for which $\overline{\rho}_{E,3}$ is reducible (or in other words, for which E admits a rational 3-isogeny), he introduced a beautiful trick, the '3-5 switch', exploiting the geometry of modular curves of low level to prove the modularity of $\overline{\rho}_{E,5}$ instead. This suffices since there are no semistable elliptic curves over $\mathbf{Q}$ with a rational 15-isogeny!

4.1 Elliptic curves over totally real fields

The strongest known modularity lifting theorem suitable for applications to the modularity of elliptic curves over totally real number fields K is the following result, taken from [19, Theorem 2]:

Theorem 4.5. *Let K be a totally real number field and let E be an elliptic curve over K. Suppose that there exists an odd prime ℓ such that the following conditions are satisfied:*

1. *$\overline{\rho}_{E,\ell}$ is modular.*
2. *$\overline{\rho}_{E,\ell}|_{G_{K(\zeta_\ell)}}$ is absolutely irreducible (here ζ_ℓ denotes a primitive ℓ^{th} root of unity in the fixed algebraic closure of K).*

Then $\rho_{E,\ell}$ is modular (and hence E itself is modular).

This is very close to optimal! The possibility of proving a theorem like this is based on numerous technical improvements to the methods introduced in [46, 41], which are

due to many people. First, one has to understand why it may be reasonable to generalise modularity lifting theorems from the case $K = \mathbf{Q}$ to the case where K is totally real. For a totally real field, the analogues of holomorphic modular forms are Hilbert modular forms. Most of the Galois representations attached to Hilbert modular forms may be constructed and analyzed using Shimura curves and the Jacquet–Langlands correspondence [12], giving a theory quite analogous to the theory of classical modular curves.

Diamond and Fujiwara [16, 21] explained how to generalise the fundamental Taylor–Wiles patching technique introduced in [41] to this context, making it possible to prove the first modularity lifting theorems over totally real fields, and also introducing soluble base change, using [31], as a fundamental tool. At this point the main question was how to impose conditions from ℓ-adic Hodge theory[3] (such as the above-mentioned weight 2 condition) while still being able to control the Galois deformation theory (in [46] only smooth conditions were considered, in which case computing the tangent space to the deformation functor in terms of Galois cohomology is enough – not so in general). This problem was solved by Kisin [29], who introduced a variant of the Taylor–Wiles method and defined and analysed weight 2 lifting functors using sophisticated results in integral ℓ-adic Hodge theory. Finally, Khare and Wintenberger, on their way to proving Serre's conjecture, introduced an important new technique for constructing liftings of modulo ℓ Galois representations with prescribed properties [28], using modularity lifting theorems and Taylor's potential automorphy technique [42] as an input. This was exploited in a very clever way by Barnet-Lamb, Gee, and Geraghty [3] in order to optimise Kisin's results.

With Theorem 4.5 in hand, we see that for an elliptic curve over a totally real field K to fail to be modular, each of its residual representations must either be degenerate (in the sense that $\overline{\rho}_{E,\ell}|_{G_{K(\zeta_\ell)}}$ is reducible) or must fail to be modular. The coincidences underlying Wiles's proof of the representations $\overline{\rho}_{E,3}$, together with the 3-5 switch, generalise well to the totally real context. Using the geometry of the modular curve $X(7)$, Manoharmayum [32] gave a 3-7 switch argument, making it possible now to prove the following theorem.

Theorem 4.6. *Let E be an elliptic curve over a totally real field K. If $\overline{\rho}_{E,\ell}|_{G_{K(\zeta_\ell)}}$ is absolutely irreducible for any of $\ell = 3, 5,$ or 7, then E is modular.*

Using this, Freitas, Le Hung and Siksek were able to prove the following striking result:

Theorem 4.7. *Let K be a totally real field. Then:*

1. *There is a finite set $S \subset K$ such that if E is an elliptic curve over K and $j(E) \notin S$, then E is modular.*
2. *If $[K : \mathbf{Q}] = 2$, then every elliptic curve over K is modular.*

[3] More normally called p-adic Hodge theory, but we consider ℓ-adic representations in this article.

(Here $j(E)$ is the j-invariant, which classifies the $\overline{K}$-isomorphism class of E.) The proof of this theorem is based on the following idea: if E is a non-modular elliptic curve then, by Theorem 4.6, it must determine a rational point on one of a finite set of modular curves parameterising elliptic curves with some of kind degeneracy of their modulo 3, 5 and 7 Galois representations. (For example, this set would include the curve $X_0(105)$, which parameterises elliptic curves for which each of the modulo 3, 5 and 7 Galois representations is reducible already on G_K.) The first part of Theorem 4.7 is thus a consequence of the observation that each of these modular curves has genus greater than 2, together with Faltings's theorem (i.e. Mordell's conjecture) [18]. The second part, much the harder, is to analyse the points of these modular curves which are defined over real quadratic fields. Similar ideas have been used by Derickx, Najman, and Siksek to establish also the modularity of elliptic curves over totally real cubic fields [15], and by Box to establish the modularity of elliptic curves over most totally real quartic fields [7].

Here is a 'vertical' analogue of Theorem 4.7 (2), proved in [44]:

Theorem 4.8. *Let p be a prime, and let $K/\mathbf{Q}$ be a totally real abelian extension, unramified away from the prime p, such that* $\mathrm{Gal}(K/\mathbf{Q})$ *has order a power of p. Then every elliptic curve over K is modular.*

This theorem is again proved by combining modularity lifting theorems and an analysis of rational points on modular curves, although in a different way. The first main ingredient is a new modularity lifting theorem, proved in [43], which removes the assumption that $\rho_{E,\ell}|_{G_{K(\zeta_\ell)}}$ is irreducible. This so-called Taylor–Wiles assumption is used to control certain Galois cohomology groups. The effect of this new theorem is that in proving Theorem 4.8, one needs consider only rational points on the single modular curve $X_0(15)$. This curve has genus 1, so could have infinitely many rational points over a fixed number field (as it does, for example, over $\mathbf{Q}(\sqrt{3})$). However, it turns out that for any field K as in the statement of Theorem 4.8, we in fact have $X_0(15)(K) = X_0(15)(\mathbf{Q})$! Any such field K is contained in the cyclotomic $\mathbf{Z}_p$-extension $\mathbf{Q}_\infty/\mathbf{Q}$, so the natural tool to prove this is Iwasawa theory, and in particular the results of Kato [27].

Looking at Theorems 4.7 and 4.8, it seems reasonable, in principle, to try to prove the modularity of all elliptic curves over any family $\mathcal{F}$ of totally real number fields for which the points of modular curves rational over members of $\mathcal{F}$ can be 'organised' in some way. Establishing the modularity of elliptic curves over all totally real fields will require new ideas.

4.2 Elliptic curves over more general number fields

We now consider the modularity of elliptic curves over number fields which are not totally real. Until a few years ago, it was very mysterious how one might hope to prove modularity lifting theorems in this context. First, it is not known in general how to associate Galois representations to Hecke eigenclasses in $H^*(Y_1(\mathfrak{n}), \mathbf{Q}_\ell)$. Indeed, the

spaces $Y_1(\mathfrak{n})$ (and their analogues, associated to quaternion algebras over number fields) have no obvious relation to algebraic geometry (for example, when K has a complex place they have no complex structure). Second, even if one could solve this problem, the spaces $Y_1(\mathfrak{n})$ can have non-trivial torsion classes in their cohomology (say with $\mathbf{Z}_\ell$ coefficients) which cannot be described in terms of automorphic representations (see e.g. [5]). Third, the Taylor–Wiles method breaks down because the cohomology groups of $Y_1(\mathfrak{n})$ (again, say, with $\mathbf{Z}_\ell$ coefficients, and now some auxiliary Taylor–Wiles level structure) are not free modules over the group rings of diamond operators that appear in the version of the Taylor–Wiles method developed by Diamond and Fujiwara.

The way forward was explained by Calegari and Geraghty [9]. Assuming a number of conjectures, they explain how to generalise the Taylor–Wiles method and prove modularity lifting theorems over general number fields which can be applied, for example, to prove the modularity of elliptic curves. We will not attempt to formulate these conjectures precisely here but note that their conjectures include the important prescription that there should exist Galois representations associated not just to (algebraic) automorphic representations with complex coefficients, but also to torsion classes in the cohomology of spaces like $Y_1(\mathfrak{n})$. This is a striking enlargement of the Langlands program as outlined in [13]!

To get unconditional results, one still has to establish the conjectures which are taken as a starting point in [9]. Progress towards these conjectures was made first by Scholze, who used his theory of perfectoid spaces to prove the existence of Galois representations attached to Hecke eigenclasses in the groups $H^*(Y_1(\mathfrak{n}), \mathbf{Z}_\ell)$ when K is a CM field, i.e. a totally imaginary quadratic extension of a totally real field [35]. Using the further results of Caraiani and Scholze on the cohomology of non-compact Shimura varieties [10], the 10 author collaboration [1] established enough of the Calegari–Geraghty conjectures to be able to establish unconditional modularity lifting theorems over CM fields. These sufficed to be able to prove, for example, the potential modularity of all elliptic curves E over CM fields K (i.e. the modularity of the base change E_L, for some finite extension L/K depending on E – a result which implies in particular the *meromorphic* continuation to $\mathbf{C}$ of $L(E, s)$).

Separately, Boxer, Calegari, Gee, and Pilloni studied the application of the Calegari–Geraghty method in the context of the coherent cohomology of Siegel type Shimura varieties [6]. The problems faced here are analogous, but different, to those arising out of the singular cohomology of the locally symmetric spaces $Y_1(\mathfrak{n})$. Nevertheless these authors were able to prove unconditional modularity lifting theorems that can be applied to the Galois representations arising from abelian surfaces over totally real fields. As a particular consequence, they are able to prove the potential modularity of elliptic curves over any quadratic extension of a totally real field (not necessarily CM) – the first general results of this kind that can be applied to elliptic curves over non-CM fields. An excellent guide to the path to the results of the last few paragraphs can be found in the survey article [11].

What about modularity (as opposed to potential modularity) of elliptic curves? To prove modularity using modularity lifting theorems, one needs a source of modular

residual representations. Unfortunately, one can no longer use Wiles's idea to prove the modularity of representations $\overline{\rho}_{E,3}$ for elliptic curves E when the base field K is not totally real. The reason is that, although the Langlands–Tunnell theorem applies over arbitrary base number fields, there is no known way to construct congruences between the automorphic representations it gives and those automorphic representations which contribute to the cohomology of locally symmetric spaces. A solution to this problem would also allow the construction of the Galois representations associated to algebraic Maass forms, a famously difficult open problem!

Nevertheless, we were able to establish the following theorem in [2]:

Theorem 4.9. *Let K be a CM field, and let E be an elliptic curve over K with multiplicative reduction at each place $v|5$ of K. Then $\overline{\rho}_{E,3}$ is modular.*

Corollary 4.10. *Let K be a CM field such that $\zeta_5 \notin K$. Then a positive proportion of elliptic curves over K are modular.*

The proof of Theorem 4.9 is based on the idea of a kind of 2-3 switch: we want to find an auxiliary elliptic curve A such that $\overline{\rho}_{A,3} \cong \overline{\rho}_{E,3}$ and $\overline{\rho}_{A,2}$ extends to a representation of G_{K^+}, where K^+ is the maximal totally real subfield of K. A tricky 2-adic modularity lifting theorem would then imply the modularity of A, hence of $\overline{\rho}_{A,3} \cong \overline{\rho}_{E,3}$. In fact, the existence of such an auxiliary curve A is a delicate matter (partly explained by the fact that the modular curve $X(6)$ has genus 1) and we need to take a more circuitous route, for which we refer to [2].

The local conditions at the 5-adic places in Theorem 4.9 are always satisfied after possibly replacing K by a soluble CM extension. Since we are free to make a soluble base change when establishing the modularity of a given elliptic curve E (by cyclic base change [31]), a sufficiently powerful modularity lifting theorem would, when combined with Theorem 4.9, prove the modularity of most elliptic curves over a given CM field.

The modularity lifting theorems established in [1] apply only to elliptic curves which have either good reduction at each place of K above the fixed prime ℓ, with ℓ unramified in K, or which have good ordinary/multiplicative reduction at each place of K above ℓ. Thus we do not have yet have access to theorems such as those proved by Kisin over totally real fields [29], in which an arbitrary amount of ramification is permitted. If such theorems can be established in the future then it seems reasonable to hope that it will be possible to prove e.g. the modularity of all elliptic curves over imaginary quadratic fields.

Acknowledgments. The author's work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 714405). I am grateful to the anonymous referee for their useful comments.

References

[1] P. B. Allen, F. Calegari, A. Caraiani, T. Gee, D. Helm, B. V. Le Hung, J. Newton, P. Scholze, R. L. Taylor, J. A. Thorne, Potential automorphy over CM fields, arXiv:1812.09999

[2] P. B. Allen, C. Khare, J. A. Thorne, Modularity of $\mathrm{GL}_2(\mathbf{F}_p)$-representations over CM fields, arXiv:1910.12986

[3] T. Barnet-Lamb, T. Gee, D. Geraghty, Congruences between Hilbert modular forms: constructing ordinary lifts. *Duke Math. J.* **161** (2012), no. 8, 1521–1580.

[4] M. A. Bennett, I. Chen, S. R. Dahmen, S. Yazdani, Generalized Fermat equations: a miscellany. *Int. J. Number Theory* **11** (2015), no. 1, 1–28.

[5] N. Bergeron, A. Venkatesh, The asymptotic growth of torsion homology for arithmetic groups. J. Inst. Math. Jussieu **12** (2013), no. 2, 391–447.

[6] G. Boxer, F. Calegari, T. Gee, V. Pilloni, Abelian surfaces over totally real fields are potentially modular, arXiv:1812.09269

[7] J. Box, Elliptic curves over totally real quartic fields not containing $\sqrt{5}$ are modular, arXiv:2103.13975

[8] C, Breuil, B. Conrad, F. Diamond, R. L. Taylor, On the modularity of elliptic curves over **Q**: wild 3-adic exercises. *J. Amer. Math. Soc.* **14** (2001), no. 4, 843–939.

[9] F. Calegari, D. Geraghty, Modularity lifting beyond the Taylor-Wiles method. *Invent. Math.* **211** (2018), no. 1, 297–433.

[10] A. Caraiani, P. Scholze, On the generic part of the cohomology of non-compact unitary Shimura varieties, arXiv:1909.01898

[11] F. Calegari, Reciprocity in the Langlands program since Fermat's Last Theorem, arXiv:2109.14145

[12] H. Carayol, Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert. *Ann. Sci. École Norm. Sup.* (4) **19** (1986), no. 3, 409–468.

[13] L. Clozel, Motifs et formes automorphes: applications du principe de fonctorialité. In *Automorphic forms, Shimura varieties, and L-functions*, Vol. I (Ann Arbor, MI, 1988), 77–159, Perspect. Math., 10, Academic Press, Boston, MA, 1990.

[14] J. E. Cremona, Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields. *Compositio Math.* **51** (1984), no. 3, 275–324.

[15] M. Derickx, F. Najman, S. Siksek, Elliptic curves over totally real cubic fields are modular. *Algebra & Number Theory* **14** (2020), no. 7, 1791–1800.

[16] F. Diamond, The Taylor–Wiles construction and multiplicity one. *Invent. Math.* **128** (1997), no. 2, 379–391.

[17] F. Diamond, J. Shurman, A first course in modular forms. Graduate Texts in Mathematics, 228. Springer-Verlag, New York, 2005.

[18] G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.* **73** (1983), no. 3, 349–366.

[19] N. Freitas, B. V. Le Hung, S. Siksek, Elliptic curves over real quadratic fields are modular. *Invent. Math.* **201** (2015), no. 1, 159–206.

[20] , N. Freitas, S. Siksek, The asymptotic Fermat's last theorem for five-sixths of real quadratic fields. *Compositio Math.* **151** (2015), no. 8, 1395–1415.

[21] K. Fujiwara, Galois deformations and arithmetic geometry of Shimura varieties. In *International Congress of Mathematicians. Vol. II*, 347–371, Eur. Math. Soc., Zürich, 2006.

[22] B. H. Gross, D. B. Zagier, Heegner points and derivatives of L-series. *Invent. Math.* **8**4 (1986), no. 2, 225–320.

[23] B. H. Gross, Lectures on the conjecture of Birch and Swinnerton-Dyer. *Arithmetic of L-functions*, 169–209, IAS/Park City Math. Ser., 18, Amer. Math. Soc., Providence, RI, 2011.

[24] F. Grunewald, H. Helling, J. Mennicke, SL_2 over complex quadratic number fields. I. *Algebra i Logika* **17** (1978), no. 5, 512–580, 622.

[25] H. Jacquet, R. P. Langlands, Automorphic forms on GL(2). *Lecture Notes in Mathematics*, Vol. 114. Springer-Verlag, Berlin-New York, 1970.

[26] R. von Känel, B. Matschke, Solving S-unit, Mordell, Thue, Thue-Mahler and generalized Ramanujan-Nagell equations via Shimura-Taniyama conjecture, arXiv:1605.06079.

[27] K. Kato, p-adic Hodge theory and values of zeta functions of modular forms. In Cohomologies p-adiques et applications arithmétiques, III. *Astérisque* No. 295 (2004), 117–290.

[28] C. Khare, J.-P. Wintenberger, On Serre's conjecture for 2-dimensional mod p representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Ann. of Math.* (2) **169** (2009), no. 1, 229–253.

[29] M. Kisin, Moduli of finite flat group schemes, and modularity. *Ann. of Math.* (2) **170** (2009), no. 3, 1085–1180.

[30] V. A. Kolyvagin, Finiteness of $E(\mathbf{Q})$ and $CH(E,\mathbf{Q})$ for a subclass of Weil curves. *Izv. Akad. Nauk SSSR Ser. Mat.* **52** (1988), no. 3, 522–540.

[31] R. P. Langlands, Base change for GL(2). Annals of Mathematics Studies, No. 96, Princeton University Press, Princeton, N.J., 1980.

[32] J. Manoharmayum, On the modularity of certain $\mathrm{GL}_2(\mathbf{F}_7)$ Galois representations. *Math. Res. Lett.* **8** (2001), 703–712.

[33] J. Newton, J. A. Thorne, Torsion Galois representations over CM fields and Hecke algebras in the derived category. Forum Math. Sigma **4** (2016), Paper No. e21, 88 pp.

[34] K. A. Ribet, On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms. *Invent. Math.* **100** (1990), no. 2, 431–476.

[35] P. Scholze, On torsion in the cohomology of locally symmetric varieties. *Ann. of Math.* (2) **182** (2015), no. 3, 945–1066.

[36] M. H. Şengün, S. Siksek, On the asymptotic Fermat's last theorem over number fields. *Comment. Math. Helv.* **93** (2018), no. 2, 359–375.

[37] J.-P. Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, *Duke Math. J.* **54** (1987), no. 1, 179–230.

[38] J. H. Silverman, The arithmetic of elliptic curves. Graduate Texts in Mathematics, 106. Springer-Verlag, New York, 1986.

[39] J. Tate, Number theoretic background. In *Automorphic forms, representations and L-functions, Part 2*, pp. 3–26, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979.

[40] R. L. Taylor, Representations of Galois groups associated to modular forms. In *Proceedings of the International Congress of Mathematicians*, Vol. 1 (Zürich, 1994), 435–442, Birkhäuser, Basel, 1995

[41] R. L. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras. *Ann. of Math.* (2) **141** (1995), no. 3, 553–572.

[42] R. L. Taylor, On the meromorphic continuation of degree two L-functions. *Doc. Math.* 2006, Extra Vol., 729–779.

[43] J. A. Thorne, Automorphy of some residually dihedral Galois representations. *Math. Ann.* **364** (2016), no. 1-2, 589–648.

[44] J. A. Thorne, Elliptic curves over $\mathbf{Q}_\infty$ are modular. *J. Eur. Math. Soc.* **21** (2019), no. 7, 1943–1948.

[45] J. Tunnell, Artin's conjecture for representations of octahedral type. *Bull. Amer. Math. Soc. (N.S.)* **5** (1981), no. 2, 173–175.

[46] A. Wiles, Modular elliptic curves and Fermat's last theorem. *Ann. of Math.* (2) **141** (1995), no. 3, 443–551.

[47] S. Zhang, Heights of Heegner points on Shimura curves. *Ann. of Math.* (2) **153** (2001), no. 1, 27–147.