WILES' PROOF OF FERMAT'S LAST THEOREM

K. RUBIN AND A. SILVERBERG

Introduction

On June 23, 1993, Andrew Wiles wrote on a blackboard, before an audience at the Newton Institute in Cambridge, England, that if p is a prime number, u, v, and w are rational numbers, and $u^p + v^p + w^p = 0$, then $uvw = 0$. In other words, he announced that he could prove Fermat's Last Theorem. His announcement came at the end of his series of three talks entitled "Modular forms, elliptic curves, and Galois representations" at the week-long workshop on "p-adic Galois representations, Iwasawa theory, and the Tamagawa numbers of motives".

In the margin of his copy of the works of Diophantus, next to a problem on Pythagorean triples, Pierre de Fermat (1601 - 1665) wrote:

> *Cubem autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caparet.*
>
> (It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second into two like powers. I have discovered a truly marvelous proof of this, which this margin is too narrow to contain.)

We restate Fermat's conjecture as follows.

Fermat's Last Theorem. *If $n > 2$, then $a^n + b^n = c^n$ has no solutions in nonzero integers a, b, and c.*

A proof by Fermat has never been found, and the problem remained open, spurring number theorists to ever greater heights. For details on the history of Fermat's Last Theorem (last because it is the last of Fermat's questions to be answered) see [5], [6], and [27].

What Andrew Wiles announced in Cambridge was that he could prove "many" elliptic curves are modular, sufficiently many to imply Fermat's Last Theorem. In this paper we will explain Wiles' result and its connection with Fermat's Last Theorem. In §1 we introduce elliptic curves and modularity,

The authors thank the NSF for financial support.

and give the connection between Fermat's Last Theorem and the Taniyama-Shimura Conjecture on the modularity of elliptic curves. In §2 we describe how Wiles reduces the proof of the Taniyama-Shimura conjecture to what we call the Modular Lifting Conjecture (which can be viewed as a weak form of the Taniyama-Shimura Conjecture), by using a theorem of Langlands and Tunnell. In §3 and §4 we show how the Modular Lifting Conjecture is related to a conjecture of Mazur on deformations of Galois representations (Conjecture 4.2), and in §5 we describe Wiles' method of attack on this conjecture. Although he does not prove the full Mazur Conjecture (and thus does not prove the full Taniyama-Shimura Conjecture), Wiles' result (Theorem 5.3) implies enough of the Modular Lifting Conjecture to prove Fermat's Last Theorem.

Much of this report is based on notes from Wiles' lectures in Cambridge. The authors apologize for any errors we may have introduced. We also apologize to those whose mathematical contributions we, due to our incomplete understanding, do not properly acknowledge.

As this paper is being completed (early November 1993), Wiles' proof is being checked by referees. Because of the great interest in this subject and the lack of a publicly available manuscript, we hope this report will be useful to the mathematics community.

In order to make this survey as accessible as possible to non-specialists, the more technical details are postponed as long as possible, some of them to the Appendices.

The integers, rational numbers, complex numbers, and p-adic integers will be denoted $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{C}$, and $\mathbf{Z}_p$, respectively. If F is a field, then $\bar{F}$ denotes an algebraic closure of F.

1. Connection between Fermat's Last Theorem and elliptic curves

1.1. Fermat's Last Theorem follows from the modularity of elliptic curves. Suppose Fermat's Last Theorem were false. Then there would exist nonzero integers a, b, c, and $n > 2$, such that $a^n + b^n = c^n$. It is easy to see that no generality is lost by assuming that n is a prime greater than three (or greater than four million, by [2]; see [14] for $n = 3$ and 4), and that a and b are relatively prime. Write down the cubic curve:

$$y^2 = x(x + a^n)(x - b^n). \tag{1}$$

In §1.3 we will see that such curves are elliptic curves, and in §1.4 we will explain what it means for an elliptic curve to be modular. Kenneth Ribet [28] proved that if n is a prime greater than three, a, b, and c are nonzero integers, and $a^n + b^n = c^n$, then the elliptic curve (1) is not modular.

Theorem 1.1 (Wiles). *If A and B are distinct, non-zero, relatively prime integers, and $AB(A-B)$ is divisible by 16, then the elliptic curve*

$$y^2 = x(x+A)(x+B)$$

is modular.

Taking $A = a^n$ and $B = -b^n$ with a, b, c, and n coming from our hypothetical solution to a Fermat equation as above, we see that the conditions of Theorem 1.1 are satisfied since $n \geq 5$ and one of a, b, and c is even. Thus Theorem 1.1 and Ribet's result together imply Fermat's Last Theorem!

1.2. History. The story of the connection between Fermat's Last Theorem and elliptic curves begins in 1955, when Yutaka Taniyama (1927 - 1958) posed problems which may be viewed as a weaker version of the following conjecture (see [39]).

Taniyama-Shimura Conjecture. *Every elliptic curve over* $\mathbf{Q}$ *is modular.*

The conjecture in the present form was made by Goro Shimura around 1962-64, and has become better understood due to work of Shimura [34], [35], [36], [37], [38] and of André Weil [43] (see also [7]).

Beginning in the late 1960's ([16], [17], [18], [19]), Yves Hellegouarch connected Fermat equations $a^n + b^n = c^n$ with elliptic curves of the form (1), and used results about Fermat's Last Theorem to prove results about elliptic curves. The landscape changed abruptly in 1985 when Gerhard Frey stated in a lecture at Oberwolfach that elliptic curves arising from counterexamples to Fermat's Last Theorem could not be modular [11]. Shortly thereafter Ribet [28] proved this, following ideas of Jean-Pierre Serre [33] (see [25] for a survey). In other words, "Taniyama-Shimura Conjecture $\Rightarrow$ Fermat's Last Theorem".

Thus, the stage was set. A proof of the Taniyama-Shimura Conjecture (or enough of it to know that elliptic curves coming from Fermat equations are modular) would be a proof of Fermat's Last Theorem.

1.3. Elliptic curves.

Definition. An *elliptic curve* over $\mathbf{Q}$ is a nonsingular curve defined by an equation of the form

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \qquad (2)$$

where the coefficients a_i are integers. The solution (∞, ∞) will be viewed as a point on the elliptic curve.

Remarks. (i) A *singular point* on a curve $f(x,y)=0$ is a point where both partial derivatives vanish. A curve is *nonsingular* if it has no singular points.
(ii) Two elliptic curves over $\mathbf{Q}$ are *isomorphic* if one can be obtained from the other by changing coordinates $x = A^2x' + B$, $y = A^3y' + Cx' + D$, with A, B, C, $D \in \mathbf{Q}$, and dividing through by A^6.
(iii) Every elliptic curve over $\mathbf{Q}$ is isomorphic to one of the form

$$y^2 = x^3 + a_2x^2 + a_4x + a_6$$

with integers a_i. A curve of this form is nonsingular if and only if the cubic on the right side has no repeated roots.

Example. The equation $y^2 = x(x+3^2)(x-4^2)$ defines an elliptic curve over $\mathbf{Q}$.

1.4. Modularity. Let $\mathfrak{H}$ denote the complex upper half plane $\{z \in \mathbf{C} : \mathrm{Im}(z) > 0\}$ where $\mathrm{Im}(z)$ is the imaginary part of z. If N is a positive integer, define a group of matrices

$$\Gamma_0(N) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbf{Z}) : c \text{ is divisible by } N \right\}.$$

The group $\Gamma_0(N)$ acts on $\mathfrak{H}$ by linear fractional transformations $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(z) = \frac{az+b}{cz+d}$. The quotient space $\mathfrak{H}/\Gamma_0(N)$ is a (non-compact) Riemann surface. It can be completed to a compact Riemann surface, denoted $X_0(N)$, by adjoining the cusps, which are the finitely many equivalence classes of $\mathbf{Q} \cup \{i\infty\}$ under the action of $\Gamma_0(N)$ (see Chapter 1 of [36]). The complex points of an elliptic curve can also be viewed as a compact Riemann surface.

Definition. An elliptic curve E is *modular* if, for some integer N, there is a holomorphic map from $X_0(N)$ onto E.

Example. There is a (holomorphic) isomorphism from $X_0(15)$ onto the elliptic curve $y^2 = x(x+3^2)(x-4^2)$.

Remark. There are many equivalent definitions of modularity (see §II.4.D of [25] and appendix of [23]). In some cases the equivalence is a deep result. For Wiles' proof of Fermat's Last Theorem it suffices to use only the definition given in §1.7 below.

1.5. Semistability.

Definition. An elliptic curve over $\mathbf{Q}$ is *semistable at the prime* q if it is isomorphic to an elliptic curve over $\mathbf{Q}$ which modulo q either is nonsingular or has a singular point with two distinct tangent directions. An elliptic curve over $\mathbf{Q}$ is called *semistable* if it is semistable at every prime.

Example. The elliptic curve $y^2 = x(x+3^2)(x-4^2)$ is semistable because it is isomorphic to $y^2 + xy + y = x^3 + x^2 - 10x - 10$, but the elliptic curve $y^2 = x(x+4^2)(x-3^2)$ is not semistable (it is not semistable at 2).

In §5.5 (Theorem 5.3) we state Wiles' main result, and explain how it implies the following theorem.

Theorem 1.2 (Wiles). *Every semistable elliptic curve over* $\mathbf{Q}$ *is modular.*

If A and B are distinct, nonzero, relatively prime integers write $E_{A,B}$ for the elliptic curve defined by $y^2 = x(x+A)(x+B)$. Since $E_{A,B}$ and $E_{-A,-B}$ are isomorphic over the complex numbers (i.e., as Riemann surfaces), $E_{A,B}$ is modular if and only if $E_{-A,-B}$ is modular. If further $AB(A-B)$ is divisible by 16, then either $E_{A,B}$ or $E_{-A,-B}$ is semistable (this is easy to check directly; see for example §I.1 of [25]), and therefore both are modular by Theorem 1.2. Thus Theorem 1.2 implies Theorem 1.1, and hence Fermat's Last Theorem.

1.6. Modular forms. In this paper we will work with a definition of modularity which uses modular forms.

Definition. If N is a positive integer, a *modular form* f of weight k for $\Gamma_0(N)$ is a holomorphic function $f : \mathfrak{H} \to \mathbf{C}$ which satisfies

$$f(\gamma(z)) = (cz+d)^k f(z) \text{ for every } \gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N) \text{ and } z \in \mathfrak{H}, \tag{3}$$

and is holomorphic at the cusps (see Chapter 2 of [36]).

A modular form f satisfies $f(z) = f(z+1)$ (apply (3) to $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$), so it has a Fourier expansion $f(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z}$, with complex numbers a_n and with $n \geq 0$ because f is holomorphic at the cusp $i\infty$. We say f is a *cusp form* if it vanishes at all the cusps; in particular for a cusp form the coefficient a_0 (the value at $i\infty$) is zero. Call a cusp form *normalized* if $a_1 = 1$.

For fixed N there are commuting linear operators (called *Hecke operators*) T_m, for integers $m \geq 1$, on the vector space of cusp forms of weight two for $\Gamma_0(N)$ (see Chapter 3 of [36]). If $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ then

$$T_m f(z) = \sum_{n=1}^{\infty} \Big(\sum_{\substack{(d,N)=1 \\ d \mid (n,m)}} d a_{nm/d^2} \Big) e^{2\pi i n z} \tag{4}$$

where (a,b) denotes the greatest common divisor of a and b and $a \mid b$ means that a divides b. The *Hecke algebra* $T(N)$ is the ring generated by these operators.

Definition. In this paper an *eigenform* will mean a normalized cusp form of weight two for some $\Gamma_0(N)$ which is an eigenfunction for all the Hecke operators.

By (4), if $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ is an eigenform then $T_m f = a_m f$ for all m.

1.7. Modularity, revisited. Suppose E is an elliptic curve over $\mathbf{Q}$. If p is a prime, write $\mathbf{F}_p$ for the finite field with p elements, and let $E(\mathbf{F}_p)$ denote the $\mathbf{F}_p$-solutions of the equation for E (including the point at infinity). We now give a second definition of modularity for an elliptic curve.

Definition. An elliptic curve E over $\mathbf{Q}$ is *modular* if there exists an eigenform $\sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ such that for all but finitely many primes q,

$$a_q = q + 1 - \#(E(\mathbf{F}_q)). \tag{5}$$

2. An overview

The flowchart (Figure 1) shows how Fermat's Last Theorem would follow if one knew the Modular Lifting Conjecture (Conjecture 2.1 below) for the primes 3 and 5. In §1.1 we discussed the upper arrow, i.e., the implication "Taniyama-Shimura Conjecture $\Rightarrow$ Fermat's Last Theorem". In this section we will discuss the other implications in the flowchart. The implication given by the lowest arrow is straightforward (Proposition 2.3), while the middle one uses an ingenious idea of Wiles (Proposition 2.4).

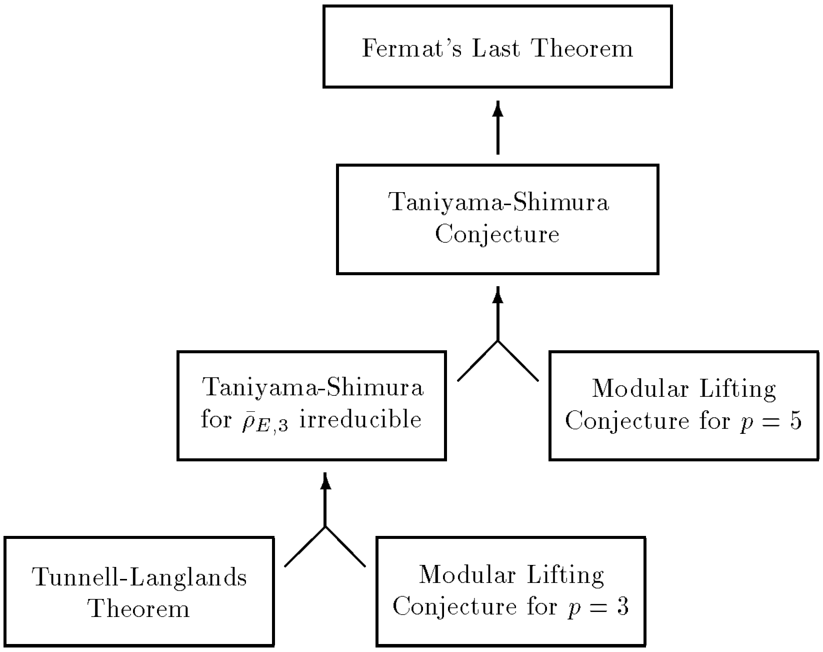

Figure 1. Modular Lifting Conjecture $\Rightarrow$ Fermat's Last Theorem

The Modular Lifting Conjecture is still an open problem, even for the primes 3 and 5. However, Wiles proves enough of the Modular Lifting Conjecture so that, with some additional work, he can still obtain enough of the Taniyama-Shimura Conjecture to prove Fermat's Last Theorem (see §5.5).

2.1. Modular Lifting Conjecture. Let $\bar{\mathbf{Q}}$ denote the algebraic closure of $\mathbf{Q}$ in $\mathbf{C}$, and let $G_{\mathbf{Q}}$ be the Galois group $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. If p is a prime, write

$$\omega_p : G_{\mathbf{Q}} \to \mathbf{F}_p^\times$$

for the character giving the action of $G_{\mathbf{Q}}$ on the p-th roots of unity. For the facts about elliptic curves stated below see [40]. If E is an elliptic curve over $\mathbf{Q}$, and F is a subfield of the complex numbers, there is a natural commutative group law on the set of F-solutions of E, with the point at infinity as the identity element. Denote this group $E(F)$. If p is a prime, write $E[p]$ for the subgroup of points in $E(\bar{\mathbf{Q}})$ of order dividing p. Then $E[p] \cong \mathbf{F}_p^2$. The action of $G_{\mathbf{Q}}$ on $E[p]$ gives a continuous representation

$$\bar{\rho}_{E,p} : G_{\mathbf{Q}} \to GL_2(\mathbf{F}_p)$$

(defined up to isomorphism) such that

$$\det(\bar{\rho}_{E,p}) = \omega_p \tag{6}$$

and for all but finitely many primes q,

$$\mathrm{trace}(\bar{\rho}_{E,p}(\mathrm{Frob}_q)) \equiv q + 1 - \#(E(\mathbf{F}_q)) \pmod{p}. \tag{7}$$

(See Appendix A for the definition of the Frobenius elements $\mathrm{Frob}_q \in G_{\mathbf{Q}}$ attached to each prime number q.)

If $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ is an eigenform, let $\mathcal{O}_f$ denote the ring of integers of the number field $\mathbf{Q}(a_2, a_3, \dots)$.

The following conjecture is in the spirit of a conjecture of Mazur (see Conjecture 3.2).

Conjecture 2.1 (Modular Lifting Conjecture). *Suppose p is a prime and E is an elliptic curve over $\mathbf{Q}$ satisfying*

(a) *$\bar{\rho}_{E,p}$ is irreducible,*
(b) *there are an eigenform $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ and a prime ideal λ of $\mathcal{O}_f$ such that $p \in \lambda$ and for all but finitely many primes q,*

$$a_q \equiv q + 1 - \#(E(\mathbf{F}_q)) \pmod{\lambda}.$$

Then E is modular.

Wiles does not prove the full Modular Lifting Conjecture, but proves it subject to some additional hypotheses on $\bar{\rho}_{E,p}$.

The Modular Lifting Conjecture is *a priori* weaker than the Taniyama-Shimura Conjecture because of the extra hypotheses (a) and (b). The more serious condition is (b); there is no known way to produce such a form in general. But when $p = 3$ the existence of such a form follows from the theorem below of Tunnell [42] and Langlands [21]. Wiles then gets around condition (a) by a clever argument (described below) which, when $\bar\rho_{E,3}$ is not irreducible, allows him to use $p = 5$ instead.

2.2. Langlands-Tunnell Theorem. In order to state the Langlands-Tunnell Theorem, we need weight one modular forms for a subgroup of $\Gamma_0(N)$. Let

$$\Gamma_1(N) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbf{Z}) : c \equiv 0 \pmod{N},\ a \equiv d \equiv 1 \pmod{N} \right\}.$$

Replacing $\Gamma_0(N)$ by $\Gamma_1(N)$ in §1.6, one can define the notion of cusp forms on $\Gamma_1(N)$. See Chapter 3 of [36] for the definitions of the Hecke operators on the space of weight one cusp forms for $\Gamma_1(N)$.

Theorem 2.2 (Langlands-Tunnell). *Suppose $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{C})$ is a continuous irreducible representation whose image in $\mathrm{PGL}_2(\mathbf{C})$ is a subgroup of S_4 (the symmetric group on four elements), τ is complex conjugation, and $\det(\rho(\tau)) = -1$. Then there is a weight one cusp form $\sum_{n=1}^{\infty} b_n e^{2\pi i n z}$ for some $\Gamma_1(N)$, which is an eigenfunction for all the corresponding Hecke operators, such that for all but finitely many primes q,*

$$b_q = \mathrm{trace}(\rho(\mathrm{Frob}_q)). \tag{8}$$

The theorem as stated by Langlands [21] and by Tunnell [42] produces an automorphic representation, rather than a cusp form. Using the fact that $\det(\rho(\tau)) = -1$, standard techniques (see for example [12]) show that this automorphic representation corresponds to a weight one cusp form as in Theorem 2.2.

2.3. Modular Lifting Conjecture $\Rightarrow$ Taniyama-Shimura Conjecture.

Proposition 2.3. *Suppose the Modular Lifting Conjecture is true for $p = 3$, E is an elliptic curve, and $\bar\rho_{E,3}$ is irreducible. Then E is modular.*

Proof. It suffices to show that hypothesis (b) of the Modular Lifting Conjecture is satisfied with the given curve E, for $p = 3$. There is a faithful representation

$$\psi : \mathrm{GL}_2(\mathbf{F}_3) \hookrightarrow \mathrm{GL}_2(\mathbf{Z}[\sqrt{-2}]) \subset \mathrm{GL}_2(\mathbf{C})$$

such that for every $g \in \mathrm{GL}_2(\mathbf{F}_3)$,

$$\mathrm{trace}(\psi(g)) \equiv \mathrm{trace}(g) \pmod{(1+\sqrt{-2})} \tag{9}$$

and

$$\det(\psi(g)) \equiv \det(g) \pmod{3}. \tag{10}$$

(Explicitly, ψ can be given by $\psi(\alpha) = \left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\psi(\beta) = \left(\begin{smallmatrix} \sqrt{-2} & 1 \\ 1 & 0 \end{smallmatrix}\right)$ where $\alpha = \left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\beta = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ generate $\mathrm{GL}_2(\mathbf{F}_3)$.) Let $\rho = \psi \circ \bar\rho_{E,3}$. If τ is complex conjugation then it follows from (6) and (10) that $\det(\rho(\tau)) = -1$.

The image of ψ in $\mathrm{PGL}_2(\mathbf{C})$ is a subgroup of $\mathrm{PGL}_2(\mathbf{F}_3) \cong S_4$. Using that $\bar{\rho}_{E,3}$ is irreducible one can show that ρ is irreducible.

Let $\mathfrak{p}$ be a prime of $\bar{\mathbf{Q}}$ containing $1+\sqrt{-2}$. Let $g(z) = \sum_{n=1}^{\infty} b_n e^{2\pi i n z}$ be a weight one cusp form for some $\Gamma_1(N)$ obtained by applying the Langlands-Tunnell Theorem (Theorem 2.2) to ρ. The function

$$\mathbf{E}(z) = 1 + 6\sum_{n=1}^{\infty}\sum_{d|n}\chi(d)e^{2\pi i n z} \quad \text{where} \quad \chi(d) = \begin{cases} 0 & \text{if } d \equiv 0 \pmod 3 \\ 1 & \text{if } d \equiv 1 \pmod 3 \\ -1 & \text{if } d \equiv 2 \pmod 3 \end{cases}$$

is a weight one modular form for $\Gamma_1(3)$. The product $g(z)\mathbf{E}(z) = \sum_{n=1}^{\infty} c_n e^{2\pi i n z}$ is a weight two cusp form for $\Gamma_0(N)$ with $c_n \equiv b_n \pmod{\mathfrak{p}}$ for all n. It is now possible to find an eigenform $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ on $\Gamma_0(N)$ such that $a_n \equiv b_n \pmod{\mathfrak{p}}$ for every n (see §6.10 of [4]). By (7), (8) and (9), f satisfies (b) of the Modular Lifting Conjecture with $p = 3$ and with $\lambda = \mathfrak{p} \cap \mathcal{O}_f$. $\square$

Proposition 2.4 (Wiles). *Suppose the Modular Lifting Conjecture is true for $p = 3$ and 5, E is an elliptic curve over $\mathbf{Q}$, and $\bar{\rho}_{E,3}$ is reducible. Then E is modular.*

Proof. The elliptic curves over $\mathbf{Q}$ for which both $\bar{\rho}_{E,3}$ and $\bar{\rho}_{E,5}$ are reducible are all known to be modular (see Appendix B.1). Thus we can suppose $\bar{\rho}_{E,5}$ is irreducible. It suffices to produce an eigenform as in (b) of the Modular Lifting Conjecture, but this time there is no analogue of the Langlands-Tunnell Theorem to help. Wiles uses the Hilbert Irreducibility Theorem, applied to a parameter space of elliptic curves, to produce another elliptic curve E' over $\mathbf{Q}$ satisfying

(i) $\bar{\rho}_{E',5}$ is isomorphic to $\bar{\rho}_{E,5}$, and
(ii) $\bar{\rho}_{E',3}$ is irreducible.

(In fact there will be infinitely many such E'; see Appendix B.2.) Now by Proposition 2.3, E' is modular. Let $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ be a corresponding eigenform. Then for all but finitely many primes q,

$$\begin{aligned} a_q = q + 1 - \#(E'(\mathbf{F}_q)) &\equiv \operatorname{trace}(\bar{\rho}_{E',5}(\mathrm{Frob}_q)) \\ &\equiv \operatorname{trace}(\bar{\rho}_{E,5}(\mathrm{Frob}_q)) \equiv q + 1 - \#(E(\mathbf{F}_q)) \pmod 5 \end{aligned}$$

by (7). Thus the form f satisfies hypothesis (b) of the Modular Lifting Conjecture and we conclude that E is modular. $\square$

Taken together Propositions 2.3 and 2.4 show that the Modular Lifting Conjecture for $p = 3$ and 5 implies the Taniyama-Shimura Conjecture.

3. Galois representations

The next step is to translate the Modular Lifting Conjecture into a conjecture (Conjecture 3.2) about the modularity of liftings of Galois representations. Throughout this paper, if A is a topological ring, a representation $\rho : G_{\mathbf{Q}} \to GL_2(A)$ will mean a continuous homomorphism and $[\rho]$ will denote the isomorphism class of ρ. If p is a prime, let

$$\varepsilon_p : G_{\mathbf{Q}} \to \mathbf{Z}_p^\times$$

be the character giving the action of $G_{\mathbf{Q}}$ on p-power roots of unity.

3.1. The p-adic representation attached to an elliptic curve. Suppose E is an elliptic curve over $\mathbf{Q}$ and p is a prime number. For every positive integer n, write $E[p^n]$ for the subgroup in $E(\bar{\mathbf{Q}})$ of points of order dividing p^n and $T_p(E)$ for the inverse limit of the $E[p^n]$ with respect to multiplication by p. For every n, $E[p^n] \cong (\mathbf{Z}/p^n\mathbf{Z})^2$, and so $T_p(E) \cong \mathbf{Z}_p^2$. The action of $G_{\mathbf{Q}}$ induces a representation

$$\rho_{E,p} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_p)$$

such that $\det(\rho_{E,p}) = \varepsilon_p$ and for all but finitely many primes q,

$$\mathrm{trace}(\rho_{E,p}(\mathrm{Frob}_q)) = q + 1 - \#(E(\mathbf{F}_q)). \tag{11}$$

Composing $\rho_{E,p}$ with the reduction map from $\mathbf{Z}_p$ to $\mathbf{F}_p$ gives $\bar{\rho}_{E,p}$ of §2.1.

3.2. Modular representations. If f is an eigenform and λ is a prime ideal of $\mathcal{O}_f$, let $\mathcal{O}_{f,\lambda}$ denote the completion of $\mathcal{O}_f$ at λ.

Definition. If A is a ring, a representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(A)$ is called *modular* if there are an eigenform $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$, a ring A' containing A, and a homomorphism $\iota : \mathcal{O}_f \to A'$ such that for all but finitely many primes q,

$$\mathrm{trace}(\rho(\mathrm{Frob}_q)) = \iota(a_q).$$

Examples. (i) Given an eigenform $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ and a prime ideal λ of $\mathcal{O}_f$, Eichler and Shimura (see §7.6 of [36]) constructed a representation

$$\rho_{f,\lambda} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathcal{O}_{f,\lambda})$$

such that $\det(\rho_{f,\lambda}) = \varepsilon_p$ (where $\lambda \cap \mathbf{Z} = p\mathbf{Z}$) and for all but finitely many primes q,

$$\mathrm{trace}(\rho_{f,\lambda}(\mathrm{Frob}_q)) = a_q. \tag{12}$$

Thus $\rho_{f,\lambda}$ is modular with ι taken to be the inclusion of $\mathcal{O}_f$ in $\mathcal{O}_{f,\lambda}$.
(ii) Suppose p is a prime and E is an elliptic curve over $\mathbf{Q}$. If E is modular, then $\rho_{E,p}$ and $\bar{\rho}_{E,p}$ are modular by (11), (7), and (5). Conversely, if $\rho_{E,p}$ is modular then it follows from (11) that E is modular. This proves the following.

Theorem 3.1. *Suppose E is an elliptic curve over $\mathbf{Q}$. Then*

E is modular $\Leftrightarrow$ $\rho_{E,p}$ is modular for every p $\Leftrightarrow$ $\rho_{E,p}$ is modular for one p.

Remark. In this language, the Modular Lifting Conjecture says that if E is an elliptic curve over $\mathbf{Q}$ and $\bar{\rho}_{E,p}$ is modular and irreducible, then $\rho_{E,p}$ is modular.

3.3. Liftings of Galois representations. Fix a prime p and a finite field k of characteristic p. Recall that $\bar{k}$ denotes an algebraic closure of k.

Given a map $\phi : A \to B$, the induced map from $\mathrm{GL}_2(A)$ to $\mathrm{GL}_2(B)$ will also be denoted ϕ.

If $\rho : G_{\mathbf{Q}} \to GL_2(A)$ is a representation and A' is a ring containing A, we write $\rho \otimes A'$ for the composition of ρ with the inclusion of $GL_2(A)$ in $GL_2(A')$.

Definition. If $\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$ is a representation, we say that a representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(A)$ is a *lifting* of $\bar{\rho}$ (to A) if A is a complete noetherian local $\mathbf{Z}_p$-algebra and there exists a homomorphism $\iota : A \to \bar{k}$ such that the diagram

$$\begin{tikzcd} & \mathrm{GL}_2(A) \arrow[d, "\iota"] \\ G_{\mathbf{Q}} \arrow[ur, "{[\rho]}"] \arrow[r, "{[\bar{\rho} \otimes \bar{k}]}"'] & \mathrm{GL}_2(\bar{k}) \end{tikzcd}$$

commutes.

Remark. Since $[\rho]$ and $[\bar{\rho}]$ are isomorphism classes of representations, the above diagram means that $\iota \circ \rho$ is isomorphic to $\bar{\rho} \otimes \bar{k}$.

Examples. (i) If E is an elliptic curve then $\rho_{E,p}$ is a lifting of $\bar{\rho}_{E,p}$.
(ii) If E is an elliptic curve, p is a prime, and hypotheses (a) and (b) of Conjecture 2.1 hold with an eigenform f and prime ideal λ, then $\rho_{f,\lambda}$ is a lifting of $\bar{\rho}_{E,p}$.

3.4. Deformation data. We will be interested not in all liftings of a given $\bar{\rho}$, but rather in those satisfying various restrictions. See Appendix A for the definition of the inertia groups $I_q \subset G_{\mathbf{Q}}$ associated to primes q. We say that a representation ρ of $G_{\mathbf{Q}}$ is *unramified* at a prime q if $\rho(I_q) = 1$. If Σ is a set of primes we say ρ is *unramified outside of* Σ if ρ is unramified at every $q \notin \Sigma$.

Definition. By *deformation data* we mean a pair

$$\mathcal{D} = (\Sigma, t)$$

where Σ is a finite set of primes and t is one of the words *ordinary* or *flat*.

If A is a $\mathbf{Z}_p$-algebra, let $\varepsilon_A : G_{\mathbf{Q}} \to \mathbf{Z}_p^\times \to A^\times$ be the composition of the cyclotomic character ε_p with the structure map.

Definition. Given deformation data $\mathcal{D}$, a representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(A)$ is *type-$\mathcal{D}$* if A is a complete noetherian local $\mathbf{Z}_p$-algebra, $\det(\rho) = \varepsilon_A$, ρ is unramified outside of Σ, and ρ is t at p (where $t \in \{\text{ordinary, flat}\}$; see Appendix C).

Definition. A representation $\bar{\rho} : G_{\mathbf{Q}} \to GL_2(k)$ is *$\mathcal{D}$-modular* if there are an eigenform f and a prime ideal λ of $\mathcal{O}_f$ such that $\rho_{f,\lambda}$ is a type-$\mathcal{D}$ lifting of $\bar{\rho}$.

Remarks. (i) A representation with a type-$\mathcal{D}$ lifting must itself be type-$\mathcal{D}$. Therefore if a representation is $\mathcal{D}$-modular then it is both type-$\mathcal{D}$ and modular. (ii) Conversely, if $\bar{\rho}$ is type-$\mathcal{D}$, modular, and satisfies (i) and (iii) of Theorem 5.3 below, then $\bar{\rho}$ is $\mathcal{D}$-modular, by work of Ribet and others (see [29]). This plays an important role in Wiles' proof of Theorem 5.3.

3.5. Mazur Conjecture.

Definition. A representation $\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$ is called *absolutely irreducible* if $\bar{\rho} \otimes \bar{k}$ is irreducible.

The following variant of a conjecture of Mazur (see Conjecture 18 of [24]; see also Conjecture 4.2 below) implies the Semistable Modular Lifting Conjecture stated below.

Conjecture 3.2 (Mazur). *Suppose p is an odd prime, k is a finite field of characteristic p, $\mathcal{D}$ is deformation data, and $\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$ is an absolutely irreducible $\mathcal{D}$-modular representation. Then every type-$\mathcal{D}$ lifting of $\bar{\rho}$ to the ring of integers of a finite extension of $\mathbf{Q}_p$ is modular.*

Remark. Loosely speaking, Conjecture 3.2 says that if $\bar{\rho}$ is modular then every lifting which "looks modular" is modular.

Conjecture 3.3 (Semistable Modular Lifting Conjecture). *Suppose p is an odd prime and E is a semistable elliptic curve over $\mathbf{Q}$ satisfying* (a) *and* (b) *of the Modular Lifting Conjecture* (*Conjecture* 2.1)*. Then E is modular.*

Proposition 3.4. *Conjecture* 3.2 *implies Conjecture* 3.3.

Proof. Suppose p is an odd prime and E is a semistable elliptic curve over $\mathbf{Q}$ which satisfies (a) and (b) of Conjecture 2.1. We will apply Conjecture 3.2 with $\bar{\rho} = \bar{\rho}_{E,p}$. Write τ for complex conjugation. Then $\tau^2 = 1$, and by (6), $\det(\bar{\rho}_{E,p}(\tau)) = -1$. Since $\bar{\rho}_{E,p}$ is irreducible and p is odd, a simple linear algebra argument now shows that $\bar{\rho}_{E,p}$ is absolutely irreducible.

Since E satisfies (b) of Conjecture 2.1, $\bar{\rho}_{E,p}$ is modular. Let

- $\Sigma = \{p\} \cup \{\text{primes } q : E \text{ has singular reduction modulo } q\}$,
- $t = $ *ordinary* if E is ordinary or singular modulo p,
  $t = $ *flat* if E is supersingular modulo p
  (see [40] for definitions of ordinary and supersingular),
- $\mathcal{D} = (\Sigma, t)$.

Using the semistability of E one can show that $\rho_{E,p}$ is a type-$\mathcal{D}$ lifting of $\bar{\rho}_{E,p}$ and (by combining results of several people; see [29]) that $\bar{\rho}_{E,p}$ is $\mathcal{D}$-modular. Conjecture 3.2 then says $\rho_{E,p}$ is modular. By Theorem 3.1, E is modular. $\square$

4. Mazur's deformation theory

Next we reformulate Conjecture 3.2 as a conjecture (Conjecture 4.2) that the algebras which parametrize liftings and modular liftings of a given representation are isomorphic. It is this form of Mazur's conjecture that Wiles attacks directly.

4.1. The universal deformation algebra R. Fix an odd prime p, a finite field k of characteristic p, deformation data $\mathcal{D}$, and an absolutely irreducible type-$\mathcal{D}$ representation $\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$. Suppose $\mathcal{O}$ is the ring of integers of a finite extension of $\mathbf{Q}_p$ with residue field k.

Definition. We say $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(A)$ is a $(\mathcal{D}, \mathcal{O})$*-lifting* of $\bar{\rho}$ if ρ is type-$\mathcal{D}$, A is a complete noetherian local $\mathcal{O}$-algebra with residue field k, and the following diagram commutes

$$\begin{tikzcd} & \mathrm{GL}_2(A) \arrow[d] \\ G_{\mathbf{Q}} \arrow[ur, "{[\rho]}"] \arrow[r, "{[\bar{\rho}]}"'] & \mathrm{GL}_2(k) \end{tikzcd}$$

where the vertical map is reduction modulo the maximal ideal of A.

Theorem 4.1 (Mazur-Ramakrishna). *With p, k, $\mathcal{D}$, $\bar{\rho}$, and $\mathcal{O}$ as above, there are an $\mathcal{O}$-algebra R and a $(\mathcal{D}, \mathcal{O})$-lifting $\rho_R : G_{\mathbf{Q}} \to \mathrm{GL}_2(R)$ of $\bar{\rho}$, with the property that for every $(\mathcal{D}, \mathcal{O})$-lifting ρ of $\bar{\rho}$ to A there is a unique $\mathcal{O}$-algebra homomorphism $\phi_\rho : R \to A$ such that the diagram*

$$\begin{tikzcd} G_{\mathbf{Q}} \arrow[r, "{[\rho_R]}"] \arrow[dr, "{[\rho]}"'] & \mathrm{GL}_2(R) \arrow[d, "\phi_\rho"] \\ & \mathrm{GL}_2(A) \end{tikzcd}$$

commutes.

This theorem was proved by Mazur [22] in the case when $\mathcal{D}$ is ordinary and by Ramakrishna [26] when $\mathcal{D}$ is flat. Theorem 4.1 determines R and ρ_R up to isomorphism.

4.2. The universal modular deformation algebra $\mathbf{T}$. Fix an odd prime p, a finite field k of characteristic p, deformation data $\mathcal{D}$, and an absolutely irreducible type-$\mathcal{D}$ representation $\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$. Assume $\bar{\rho}$ is $\mathcal{D}$-modular, and fix an eigenform f and a prime ideal λ of $\mathcal{O}_f$ such that $\rho_{f,\lambda}$ is a type-$\mathcal{D}$ lifting of $\bar{\rho}$. Suppose in addition that $\mathcal{O}$ is the ring of integers of a finite extension of $\mathbf{Q}_p$ with residue field k, $\mathcal{O}_{f,\lambda} \subseteq \mathcal{O}$, and the diagram

$$\begin{tikzcd} & \mathrm{GL}_2(\mathcal{O}_{f,\lambda}) \arrow[d] \\ G_{\mathbf{Q}} \arrow[ur, "{[\rho_{f,\lambda}]}"] \arrow[r, "{[\bar\rho]}"'] & \mathrm{GL}_2(k) \end{tikzcd}$$

commutes, where the vertical map is the reduction map.

Under these assumptions $\rho_{f,\lambda} \otimes \mathcal{O}$ is a $(\mathcal{D}, \mathcal{O})$-lifting of $\bar\rho$, and Wiles constructs a generalized Hecke algebra $\mathbf{T}$ which has the following properties.

($\mathbf{T}1$) $\mathbf{T}$ is a complete noetherian local $\mathcal{O}$-algebra with residue field k.

($\mathbf{T}2$) There are an integer N divisible only by primes in Σ and a homomorphism from the Hecke algebra $T(N)$ to $\mathbf{T}$ such that $\mathbf{T}$ is generated over $\mathcal{O}$ by the images of the Hecke operators T_q for primes $q \notin \Sigma$. By abuse of notation we write T_q also for its image in $\mathbf{T}$.

($\mathbf{T}3$) There is a $(\mathcal{D}, \mathcal{O})$-lifting

$$\rho_{\mathbf{T}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{T})$$

of $\bar\rho$ with the property that $\mathrm{trace}(\rho_{\mathbf{T}}(\mathrm{Frob}_q)) = T_q$ for every prime $q \notin \Sigma$.

($\mathbf{T}4$) If ρ is modular and is a $(\mathcal{D}, \mathcal{O})$-lifting of $\bar\rho$ to A, then there is a unique $\mathcal{O}$-algebra homomorphism $\psi_\rho : \mathbf{T} \to A$ such that the diagram

$$\begin{tikzcd} G_{\mathbf{Q}} \arrow[r, "{[\rho_{\mathbf{T}}]}"] \arrow[dr, "{[\rho]}"'] & \mathrm{GL}_2(\mathbf{T}) \arrow[d, "\psi_\rho"] \\ & \mathrm{GL}_2(A) \end{tikzcd}$$

commutes.

Since $\rho_{\mathbf{T}}$ is a $(\mathcal{D}, \mathcal{O})$-lifting of $\bar\rho$, by Theorem 4.1 there is a homomorphism

$$\varphi : R \to \mathbf{T}$$

such that $\rho_{\mathbf{T}}$ is isomorphic to $\varphi \circ \rho_R$. By ($\mathbf{T}3$), $\varphi(\mathrm{trace}(\rho_R(\mathrm{Frob}_q))) = T_q$ for every prime $q \notin \Sigma$, so it follows from ($\mathbf{T}2$) that φ is surjective.

4.3. Mazur Conjecture, revisited. Conjecture 3.2 can be reformulated in the following way.

Conjecture 4.2 (Mazur). *Suppose p, k, $\mathcal{D}$, $\bar\rho$, and $\mathcal{O}$ are as in §4.2. Then the above map $\varphi : R \to \mathbf{T}$ is an isomorphism.*

Conjecture 4.2 was stated in [24] (Conjecture 18) for $\mathcal{D}$ ordinary, and was extended to the flat case by Wiles.

Proposition 4.3. *Conjecture 4.2 implies Conjecture 3.2.*

Proof. Suppose $\bar\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(k)$ is absolutely irreducible and $\mathcal{D}$-modular, A is the ring of integers of a finite extension of $\mathbf{Q}_p$, and ρ is a type-$\mathcal{D}$ lifting of $\bar\rho$ to A. Taking $\mathcal{O}$ to be the ring of integers of a sufficiently large finite extension of $\mathbf{Q}_p$, and extending ρ and $\bar\rho$ to $\mathcal{O}$ and its residue field, respectively, we may assume that ρ is a $(\mathcal{D}, \mathcal{O})$-lifting of $\bar{\rho}$. Assuming Conjecture 4.2, let $\psi = \phi_\rho \circ \varphi^{-1} : \mathbf{T} \to A$, with ϕ_ρ as in Theorem 4.1. By (**T**3) and Theorem 4.1, $\psi(T_q) = \mathrm{trace}(\rho(\mathrm{Frob}_q))$ for all but finitely many q. By §3.5 of [36], given such a homomorphism ψ (and viewing A as a subring of $\mathbf{C}$) there is an eigenform $\sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ where $a_q = \psi(T_q)$ for all but finitely many primes q. Thus ρ is modular. □

5. Wiles' proof of part of the Mazur Conjecture

In this section we sketch the major ideas of Wiles' proof of a large part of Conjecture 4.2. The first step (Theorem 5.2), and the key to Wiles' proof, is to reduce Conjecture 4.2 to a bound on the order of the cotangent space at a prime of R. In §5.2 we see that the corresponding tangent space is a Selmer group, and in §5.3 we outline a general procedure due to Kolyvagin for bounding sizes of Selmer groups. The input for Kolyvagin's method is known as an Euler system. The most difficult part of Wiles' proof (§5.4) is his construction of a suitable Euler system. In §5.5 we state Wiles' result and explain why it suffices for proving Theorem 1.2.

For §5 fix p, k, $\mathcal{D}$, $\bar{\rho}$, $\mathcal{O}$, $f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$, and λ as in §4.2.

By property (**T**4) there is a homomorphism

$$\pi : \mathbf{T} \to \mathcal{O}$$

such that $\pi \circ \rho_{\mathbf{T}}$ is isomorphic to $\rho_{f,\lambda} \otimes \mathcal{O}$. By property (**T**2) and (12), π satisfies $\pi(T_q) = a_q$ for all but finitely many q.

5.1. Key reduction. Generalizing a result of Mazur, Wiles proves the following (Gorenstein) property of $\mathbf{T}$.

Theorem 5.1. *There is a (non-canonical)* $\mathbf{T}$*-module isomorphism*

$$\mathrm{Hom}_{\mathcal{O}}(\mathbf{T}, \mathcal{O}) \xrightarrow{\sim} \mathbf{T}.$$

Let η denote the ideal of $\mathcal{O}$ generated by the image of the element $\pi \in \mathrm{Hom}_{\mathcal{O}}(\mathbf{T}, \mathcal{O})$ under the composition

$$\mathrm{Hom}_{\mathcal{O}}(\mathbf{T}, \mathcal{O}) \xrightarrow{\sim} \mathbf{T} \xrightarrow{\pi} \mathcal{O}.$$

The ideal η is well-defined independent of the choice of isomorphism in Theorem 5.1.

The map π determines distinguished prime ideals of $\mathbf{T}$ and R,

$$\mathfrak{p}_{\mathbf{T}} = \ker(\pi), \qquad \mathfrak{p}_R = \ker(\pi \circ \varphi) = \varphi^{-1}(\mathfrak{p}_{\mathbf{T}}).$$

Theorem 5.2 (Wiles). *If*

$$\#(\mathfrak{p}_R/\mathfrak{p}_R^2) \le \#(\mathcal{O}/\eta) < \infty$$

then $\varphi : R \to \mathbf{T}$ *is an isomorphism.*

The proof is entirely commutative algebra. The surjectivity of φ shows that $\#(\mathfrak{p}_R/\mathfrak{p}_R^2) \geq \#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)$, and Wiles proves $\#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2) \geq \#(\mathcal{O}/\eta)$. Thus if $\#(\mathfrak{p}_R/\mathfrak{p}_R^2) \leq \#(\mathcal{O}/\eta)$ then

$$\#(\mathfrak{p}_R/\mathfrak{p}_R^2) = \#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2) = \#(\mathcal{O}/\eta). \tag{13}$$

The first equality in (13) shows that φ induces an isomorphism of tangent spaces. Wiles uses the second equality in (13) and Theorem 5.1 to deduce that $\mathbf{T}$ is a local complete intersection over $\mathcal{O}$ (see [15] for the definition). Wiles then combines these two results to prove that φ is an isomorphism.

5.2. Selmer groups. In general, if M is a torsion $G_{\mathbf{Q}}$-module, a Selmer group attached to M is a subgroup of the Galois cohomology group $H^1(G_{\mathbf{Q}}, M)$ determined by certain "local conditions" in the following way. If q is a prime with decomposition group $D_q \subset G_{\mathbf{Q}}$, then there is a restriction map

$$\mathrm{res}_q : H^1(G_{\mathbf{Q}}, M) \to H^1(D_q, M).$$

For a fixed collection of subgroups $\mathcal{J} = \{J_q \subseteq H^1(D_q, M) : q \text{ prime}\}$ depending on the particular problem under consideration, the corresponding Selmer group is

$$S(M) = \bigcap_q \mathrm{res}_q^{-1}(J_q) \subseteq H^1(G_{\mathbf{Q}}, M).$$

Write $H^i(\mathbf{Q}, M)$ for $H^i(G_{\mathbf{Q}}, M)$, and $H^i(\mathbf{Q}_q, M)$ for $H^i(D_q, M)$.

Example. The original examples of Selmer groups come from elliptic curves. Fix an elliptic curve E and a positive integer m, and take $M = E[m]$, the subgroup of points in $E(\bar{\mathbf{Q}})$ of order dividing m. There is a natural inclusion

$$E(\mathbf{Q})/mE(\mathbf{Q}) \hookrightarrow H^1(\mathbf{Q}, E[m]) \tag{14}$$

obtained by sending $x \in E(\mathbf{Q})$ to the cocycle $\sigma \mapsto \sigma(y) - y$, where $y \in E(\bar{\mathbf{Q}})$ is any point satisfying $my = x$. Similarly, for every prime q there is a natural inclusion

$$E(\mathbf{Q}_q)/mE(\mathbf{Q}_q) \hookrightarrow H^1(\mathbf{Q}_q, E[m]).$$

Define the Selmer group $S(E[m])$ in this case by taking the group J_q to be the image of $E(\mathbf{Q}_q)/mE(\mathbf{Q}_q)$ in $H^1(\mathbf{Q}_q, E[m])$, for every q. This Selmer group is an important tool in studying the arithmetic of E because it contains (via (14)) $E(\mathbf{Q})/mE(\mathbf{Q})$.

Retaining the notation from the beginning of §5, let $\mathfrak{m}$ denote the maximal ideal of $\mathcal{O}$ and fix a positive integer n. The tangent space $\mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, \mathcal{O}/\mathfrak{m}^n)$ can be identified with a Selmer group as follows.

Let V_n be the matrix algebra $\mathrm{M}_2(\mathcal{O}/\mathfrak{m}^n)$, with $G_{\mathbf{Q}}$ acting via the adjoint representation $\sigma(B) = \rho_{f,\lambda}(\sigma) B \rho_{f,\lambda}(\sigma)^{-1}$. There is a natural injection

$$s : \mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, \mathcal{O}/\mathfrak{m}^n) \hookrightarrow H^1(\mathbf{Q}, V_n)$$

which is described in Appendix D (see also §1.6 of [22]). Wiles defines a collection $\mathcal{J} = \{J_q \subseteq H^1(\mathbf{Q}_q, V_n)\}$ depending on $\mathcal{D}$. Let $S_{\mathcal{D}}(V_n)$ denote the associated Selmer group. Wiles proves that s induces an isomorphism

$$\mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, \mathcal{O}/\mathfrak{m}^n) \xrightarrow{\sim} S_{\mathcal{D}}(V_n).$$

5.3. Euler systems. We have now reduced the proof of Mazur's conjecture to bounding the size of the Selmer groups $S_{\mathcal{D}}(V_n)$. About five years ago Kolyvagin [20], building on ideas of his own and of Thaine [41], introduced a revolutionary new method for bounding the size of a Selmer group. This new machinery, which is crucial for Wiles' proof, is what we now describe.

Suppose M is a $G_{\mathbf{Q}}$-module of odd exponent m and $\mathcal{J} = \{J_q \subseteq H^1(\mathbf{Q}_q, M)\}$ is a system of subgroups with associated Selmer group $S(M)$ as in §5.2. Let $\hat{M} = \mathrm{Hom}(M, \boldsymbol{\mu}_m)$, where $\boldsymbol{\mu}_m$ is the group of m-th roots of unity. For every prime q, the cup product gives a nondegenerate Tate pairing

$$\langle\ ,\ \rangle_q : H^1(\mathbf{Q}_q, M) \times H^1(\mathbf{Q}_q, \hat{M}) \to H^2(\mathbf{Q}_q, \boldsymbol{\mu}_m) \xrightarrow{\sim} \mathbf{Z}/m\mathbf{Z}$$

(see Chapters VI and VII of [3]). If $c \in H^1(\mathbf{Q}, M)$ and $d \in H^1(\mathbf{Q}, \hat{M})$, then

$$\sum_q \langle \mathrm{res}_q(c), \mathrm{res}_q(d)\rangle_q = 0. \tag{15}$$

Suppose that $\mathcal{L}$ is a finite set of primes. Let $S^*_{\mathcal{L}} \subseteq H^1(\mathbf{Q}, \hat{M})$ be the Selmer group given by the local conditions $\mathcal{J}^* = \{J^*_q \subseteq H^1(\mathbf{Q}_q, \hat{M})\}$, where

$$J^*_q = \begin{cases} \text{the orthogonal complement of } J_q \text{ under } \langle\ ,\ \rangle_q & \text{if } q \notin \mathcal{L} \\ H^1(\mathbf{Q}_q, \hat{M}) & \text{if } q \in \mathcal{L}. \end{cases}$$

If $d \in H^1(\mathbf{Q}, \hat{M})$ define

$$\theta_d : \prod_{q \in \mathcal{L}} J_q \to \mathbf{Z}/m\mathbf{Z}$$

by

$$\theta_d((c_q)) = \sum_{q \in \mathcal{L}} \langle c_q, \mathrm{res}_q(d)\rangle_q.$$

Write $\mathrm{res}_{\mathcal{L}} : H^1(\mathbf{Q}, M) \to \prod_{q \in \mathcal{L}} H^1(\mathbf{Q}_q, M)$ for the product of the restriction maps. By (15) and the definition of J^*_q, if $d \in S^*_{\mathcal{L}}$ then $\mathrm{res}_{\mathcal{L}}(S(M)) \subseteq \ker(\theta_d)$. If in addition $\mathrm{res}_{\mathcal{L}}$ is injective on $S(M)$ then

$$\#(S(M)) \leq \#\big(\bigcap_{d \in S^*_{\mathcal{L}}} \ker(\theta_d)\big).$$

The difficulty is to produce enough cohomology classes in $S^*_{\mathcal{L}}$ to show that the right side of the above inequality is small. Following Kolyvagin, an Euler system is a compatible collection of classes $\kappa(\mathcal{L}) \in S^*_{\mathcal{L}}$ for a large (infinite) collection of sets of primes $\mathcal{L}$. Loosely speaking, compatible means that if

$\ell \notin \mathcal{L}$, then $\kappa(\mathcal{L} \cup \{\ell\})$ is related to $\mathrm{res}_\ell(\kappa(\mathcal{L}))$. Once an Euler system is given, Kolyvagin has an inductive procedure for choosing a set $\mathcal{L}$ such that

- $res_{\mathcal{L}}$ is injective on $S(M)$,
- $\bigcap_{\mathcal{P} \subseteq \mathcal{L}} \ker(\theta_{\kappa(\mathcal{P})})$ can be computed in terms of $\kappa(\emptyset)$.

(Note that if $\mathcal{P} \subseteq \mathcal{L}$ then $S^*_{\mathcal{P}} \subseteq S^*_{\mathcal{L}}$ so $\kappa(\mathcal{P}) \in S^*_{\mathcal{L}}$.)

For several important Selmer groups (including the one defined by Wiles) it is possible to construct Euler systems for which Kolyvagin's procedure produces a set $\mathcal{L}$ actually giving an equality

$$\#(S(M)) = \#\big(\bigcap_{\mathcal{P} \subseteq \mathcal{L}} \ker(\theta_{\kappa(\mathcal{P})})\big).$$

There are several examples in the literature where this kind of argument is worked out in some detail. For the simplest case, where the Selmer group in question is the ideal class group of a real abelian number field and the $\kappa(\mathcal{L})$ are constructed from cyclotomic units, see [30]. For other cases involving ideal class groups and Selmer groups of elliptic curves, see [20], [32], [31], [13].

5.4. Wiles' geometric Euler system. The task now is to construct an Euler system of cohomology classes with which to bound $\#(S_{\mathcal{D}}(V_n))$ using Kolyvagin's method. This is the most technically difficult part of Wiles' proof.

Unfortunately, space considerations (and our incomplete understanding) make it impossible to give the details of Wiles' truly marvelous construction.

The first step in the construction is due to Flach [10]. He constructed classes $\kappa(\mathcal{L}) \in S^*_{\mathcal{L}}$ for sets $\mathcal{L}$ consisting of just one prime. This allows one to bound the exponent of $S_{\mathcal{D}}(V_n)$, but not its order.

Every Euler system starts with some explicit, concrete objects. Earlier examples of Euler systems come from cyclotomic or elliptic units, Gauss sums, or Heegner points on elliptic curves. Wiles (following Flach) constructs his cohomology classes from modular units, i.e., meromorphic functions on modular curves which are holomorphic and nonzero away from the cusps. More precisely, $\kappa(\mathcal{L})$ comes from an explicit function on the modular curve $X_1(L, N)$, the curve obtained by taking the quotient space of the upper half plane by the action of the group

$$\Gamma_1(L,N) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z}) : c \equiv 0 \pmod{LN}, \quad a \equiv d \equiv 1 \pmod{L}\},$$

and adjoining the cusps, where $L = \prod_{\ell \in \mathcal{L}} \ell$ and where N is the N of $(\mathbf{T}2)$ of §4.2. The construction and study of the classes $\kappa(\mathcal{L})$ rely heavily on results of Faltings [8], [9] and others.

5.5. Final result. In the end, Wiles is only able to construct an Euler system and prove the desired inequality under some extra hypotheses. Since for ease of exposition we defined modularity of representations in terms of $\Gamma_0(N)$ instead of $\Gamma_1(N)$, the theorem stated below is weaker than that stated by Wiles, but has the same applications to elliptic curves.

Recall that ω_p is the character giving the action of $G_{\mathbf{Q}}$ on $\boldsymbol{\mu}_p$. If $\bar{\rho}$ is a representation of $G_{\mathbf{Q}}$ on a vector space V, $\mathrm{Sym}^2(\bar{\rho})$ denotes the representation on the symmetric square of V induced by $\bar{\rho}$.

Theorem 5.3 (Wiles). *Suppose p, k, $\mathcal{D}$, $\bar{\rho}$, and $\mathcal{O}$ are as in §4.2 and $\bar{\rho}$ satisfies the following additional conditions:*

(i) $\det(\bar{\rho}) = \omega_p$,
(ii) $\mathrm{Sym}^2(\bar{\rho})$ *is absolutely irreducible,*
(iii) *if $\bar{\rho}$ is ramified at q and $q \neq p$ then the restriction of $\bar{\rho}$ to D_q is reducible,*
(iv) *if p is 3 or 5 then for some prime q, p divides $\#(\bar{\rho}(I_q))$.*

Then $\varphi : R \to \mathbf{T}$ is an isomorphism.

Since Wiles does not prove the full Mazur Conjecture (Conjecture 4.2) for $p = 3$ and 5, we need to reexamine the arguments of §2 to see which elliptic curves E can be proved modular using Theorem 5.3 applied to $\bar{\rho}_{E,3}$ and $\bar{\rho}_{E,5}$.

By (6), hypothesis (i) of Theorem 5.3 is satisfied for every $\bar{\rho}_{E,p}$. Hypothesis (ii) will be satisfied if the image of $\bar{\rho}_{E,p}$ is sufficiently large in $\mathrm{GL}_2(\mathbf{F}_p)$ (for example, if $\bar{\rho}_{E,p}$ is surjective). For $p = 3$ and $p = 5$, if $\bar{\rho}_{E,p}$ satisfies hypothesis (iv) and is irreducible then it satisfies hypothesis (ii).

If E is semistable, p is an odd prime, and $\bar{\rho}_{E,p}$ is irreducible and modular, then $\bar{\rho}_{E,p}$ is $\mathcal{D}$-modular for some $\mathcal{D}$ (see the proof of Proposition 3.4) and $\bar{\rho}_{E,p}$ satisfies (iii) and (iv) (use Tate curves; see §14, Appendix C of [40]). Therefore by Propositions 4.3 and 3.4, Theorem 5.3 implies that the Semistable Modular Lifting Conjecture (Conjecture 3.3) holds for $p = 3$ and for $p = 5$.

Using the Langlands-Tunnell Theorem (Theorem 2.2), the same arguments that proved Propositions 2.3 and 2.4 can now be used to prove that every semistable elliptic curve over $\mathbf{Q}$ is modular (Theorem 1.2). But now the elliptic curve E' satisfying (i) and (ii) of the proof of Proposition 2.4 should also be chosen 5-adically close to E so that it will be semistable.

Appendix A. Galois groups and Frobenius elements

Write $G_{\mathbf{Q}} = \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. If q is a prime number and $\mathcal{Q}$ is a prime ideal dividing q in the ring of integers of $\bar{\mathbf{Q}}$, there is a filtration

$$G_{\mathbf{Q}} \supset D_{\mathcal{Q}} \supset I_{\mathcal{Q}}$$

where the decomposition group $D_{\mathcal{Q}}$ and the inertia group $I_{\mathcal{Q}}$ are defined by

$$\begin{aligned} D_{\mathcal{Q}} &= \{\sigma \in G_{\mathbf{Q}} : \sigma\mathcal{Q} = \mathcal{Q}\} \\ I_{\mathcal{Q}} &= \{\sigma \in \mathcal{D}_{\mathcal{Q}} : \sigma x \equiv x \pmod{\mathcal{Q}} \text{ for all algebraic integers } x\}. \end{aligned}$$

There are natural identifications

$$D_{\mathcal{Q}} \cong \mathrm{Gal}(\bar{\mathbf{Q}}_q/\mathbf{Q}_q), \qquad D_{\mathcal{Q}}/I_{\mathcal{Q}} \cong \mathrm{Gal}(\bar{\mathbf{F}}_q/\mathbf{F}_q),$$

and $\mathrm{Frob}_{\mathcal{Q}} \in D_{\mathcal{Q}}/I_{\mathcal{Q}}$ denotes the inverse image of the canonical generator $x \mapsto x^q$ of $\mathrm{Gal}(\bar{\mathbf{F}}_q/\mathbf{F}_q)$. If $\mathcal{Q}'$ is another prime ideal above q, then $\mathcal{Q}' = \sigma\mathcal{Q}$ for some $\sigma \in G_{\mathbf{Q}}$ and

$$D_{\mathcal{Q}'} = \sigma D_{\mathcal{Q}} \sigma^{-1}, \qquad I_{\mathcal{Q}'} = \sigma I_{\mathcal{Q}} \sigma^{-1}, \qquad \mathrm{Frob}_{\mathcal{Q}'} = \sigma \mathrm{Frob}_{\mathcal{Q}} \sigma^{-1}.$$

Since we will care about these objects only up to conjugation, we will write D_q and I_q. We will write $\mathrm{Frob}_q \in G_{\mathbf{Q}}$ for any representative of a $\mathrm{Frob}_{\mathcal{Q}}$. If ρ is a representation of $G_{\mathbf{Q}}$ which is unramified at q then $\mathrm{trace}(\rho(\mathrm{Frob}_q))$ and $\det(\rho(\mathrm{Frob}_q))$ are well-defined independent of any choices.

Appendix B. Some details on the proof of Proposition 2.4

B.1. The modular curve $X_0(15)$ can be viewed as a curve defined over $\mathbf{Q}$ in such a way that the non-cusp rational points correspond to isomorphism classes (over $\mathbf{C}$) of pairs $(E', \mathcal{C})$ where E' is an elliptic curve over $\mathbf{Q}$ and $\mathcal{C} \subset E(\bar{\mathbf{Q}})$ is a subgroup of order 15 stable under $G_{\mathbf{Q}}$. An equation for $X_0(15)$ is $y^2 = x(x+3^2)(x-4^2)$, the elliptic curve discussed in §1. There are eight rational points on $X_0(15)$, four of which are cusps. There are four modular elliptic curves, corresponding to a modular form for $\Gamma_0(50)$ (see p. 86 of [1]), which lie in the four distinct $\mathbf{C}$-isomorphism classes that correspond to the non-cusp rational points on $X_0(15)$.

Therefore every elliptic curve over $\mathbf{Q}$ with a $G_{\mathbf{Q}}$-stable subgroup of order 15 is modular. Equivalently, if E is an elliptic curve over $\mathbf{Q}$ and both $\bar{\rho}_{E,3}$ and $\bar{\rho}_{E,5}$ are reducible, then E is modular.

B.2. Fix an elliptic curve E over $\mathbf{Q}$. We will show that there are infinitely many elliptic curves E' over $\mathbf{Q}$ such that

(i) $\bar{\rho}_{E',5}$ is isomorphic to $\bar{\rho}_{E,5}$, and
(ii) $\bar{\rho}_{E',3}$ is irreducible.

Let

$$\Gamma(5) = \{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z}) : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \pmod 5 \}.$$

Let X be the twist of the classical modular curve $X(5)$ (see [36]) by the cocycle induced by $\bar{\rho}_{E,5}$, and let S be the set of cusps of X. Then X is a curve defined over $\mathbf{Q}$ which has the following properties.

- The rational points on $X - S$ correspond to isomorphism classes of pairs (E', ϕ) where E' is an elliptic curve over $\mathbf{Q}$ and $\phi : E[5] \to E'[5]$ is a $G_{\mathbf{Q}}$-module isomorphism.
- As a complex manifold $X - S$ is four copies of $\mathfrak{H}/\Gamma(5)$, so each component of X has genus zero.

A curve of genus zero has infinitely many rational points if it has any. Since X has a rational point corresponding to $(E, \text{identity})$, one of the components X^0 of X is a curve defined over $\mathbf{Q}$ which has infinitely many rational points. We want to show that infinitely many of these points correspond to elliptic curves E' with $\bar{\rho}_{E',3}$ irreducible.

There is another modular curve $\hat{X}$ defined over $\mathbf{Q}$, with a finite set $\hat{S}$ of cusps, which has the following properties.

- The rational points on $\hat{X} - \hat{S}$ correspond to isomorphism classes of triples $(E', \phi, \mathcal{C})$ where E' is an elliptic curve over $\mathbf{Q}$, $\phi : E[5] \to E'[5]$ is a $G_{\mathbf{Q}}$-module isomorphism, and $\mathcal{C} \subset E'[3]$ is a $G_{\mathbf{Q}}$-stable subgroup of order 3.
- As a complex manifold $\hat{X} - \hat{S}$ is four copies of $\mathfrak{H}/(\Gamma(5) \cap \Gamma_0(3))$.
- The map that forgets the subgroup $\mathcal{C}$ induces a surjective morphism $\theta : \hat{X} \to X$ defined over $\mathbf{Q}$ and of degree $[\Gamma(5) : \Gamma(5) \cap \Gamma_0(3)] = 4$.

Let $\hat{X}^0$ be the component of $\hat{X}$ which maps to X^0. The function field of X^0 is $\mathbf{Q}(t)$, and the function field of $\hat{X}^0$ is $\mathbf{Q}(t)[x]/f(t,x)$ where $f(t,x) \in \mathbf{Q}(t)[x]$ is irreducible and has degree 4 in x. By the Hilbert Irreducibility Theorem, there are infinitely many values $t' \in \mathbf{Q}$ for which $f(t', x)$ is irreducible in $\mathbf{Q}[x]$. Each such t' corresponds to a rational point of X^0 which is not the image of a rational point of $\hat{X}^0$. In other words, there are infinitely many elliptic curves E' over $\mathbf{Q}$ such that

(i) $E'[5] \cong E[5]$ as $G_{\mathbf{Q}}$-modules, and
(ii) $E'[3]$ has no subgroup of order 3 stable under $G_{\mathbf{Q}}$.

These are precisely the desired conditions on E'.

Appendix C. Representation types

Suppose A is a complete noetherian local $\mathbf{Z}_p$-algebra and $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(A)$ is a representation. Write $\rho\mid_{D_p}$ for the restriction of ρ to the decomposition group D_p. We say ρ is

- *ordinary* at p if $\rho\mid_{D_p}$ is (after a change of basis, if necessary) of the form $\left(\begin{smallmatrix} * & * \\ 0 & \chi \end{smallmatrix}\right)$ where χ is unramified and the $*$ are functions from D_p to A.
- *flat* at p if ρ is not ordinary, and for every ideal $\mathfrak{a}$ of finite index in A, the reduction of $\rho\mid_{D_p}$ modulo $\mathfrak{a}$ is the representation associated to the $\bar{\mathbf{Q}}_p$-points of a finite flat group scheme over $\mathbf{Z}_p$.

Appendix D. Selmer groups

With notation as in §5 (see especially §5.2), define

$$\mathcal{O}_n = \mathcal{O}[\epsilon]/(\epsilon^2, \mathfrak{m}^n)$$

where ϵ is an indeterminate. Then $v \mapsto 1 + \epsilon v$ defines an isomorphism

$$V_n \xrightarrow{\sim} \{\delta \in \mathrm{GL}_2(\mathcal{O}_n) : \delta \equiv 1 \pmod{\epsilon}\}. \tag{16}$$

For every $\alpha \in \mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, \mathcal{O}/\mathfrak{m}^n)$ there is a unique $\mathcal{O}$-algebra homomorphism $\psi_\alpha : R \to \mathcal{O}_n$ whose restriction to $\mathfrak{p}_R$ is $\epsilon\alpha$. Composing with the representation ρ_R of Theorem 4.1 gives a $(\mathcal{D}, \mathcal{O})$-lifting $\rho_\alpha = \psi_\alpha \circ \rho_R$ of $\bar{\rho}$ to $\mathcal{O}_n$. (In particular ρ_0 denotes the $(\mathcal{D}, \mathcal{O})$-lifting obtained when $\alpha = 0$.) Define a one-cocycle c_α on $G_{\mathbf{Q}}$ by

$$c_\alpha(g) = \rho_\alpha(g)\rho_0(g)^{-1}.$$

Since $\rho_\alpha \equiv \rho_0 \pmod{\epsilon}$, using (16) we can view $c_\alpha \in H^1(\mathbf{Q}, V_n)$. This defines a homomorphism

$$s : \mathrm{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, \mathcal{O}/\mathfrak{m}^n) \to H^1(\mathbf{Q}, V_n)$$

and it is not difficult to see that s is injective. The fact that ρ_0 and ρ_α are type-$\mathcal{D}$ gives information about the restrictions $\mathrm{res}_q(c_\alpha)$ for various primes q, and using this information Wiles defines a Selmer group $S_{\mathcal{D}}(V_n) \subset H^1(\mathbf{Q}, V_n)$ and verifies that s is an isomorphism onto $S_{\mathcal{D}}(V_n)$.

References

1. Birch, B., Kuyk, W., eds., Modular functions of one variable IV, vol. 476, Springer-Verlag, New York (1975) 74–144.
2. Buhler, J., Crandall, R., Ernvall, R., Metsänkylä, T., *Irregular Primes and Cyclotomic Invariants to Four Million*, Math. Comp. 61 (1993) 151-153.
3. Cassels, J.W.S., Frohlich, A., Algebraic number theory, Academic Press, London (1967).
4. Deligne, P., Serre, J-P., *Formes modulaires de poids* 1, Ann. Sci. Ec. Norm. Sup. **7** (1974) 507–530.
5. Dickson, L. E., History of the theory of numbers (Vol. II), Chelsea Publ. Co., New York (1971).
6. Edwards, H. M., Fermat's Last Theorem. A genetic introduction to algebraic number theory. Springer-Verlag, New York (1977).
7. Eichler, M., *Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion*, Arch. Math. **5** (1954) 355–366.
8. Faltings, G., *p-adic Hodge theory*, J. of the A. M. S. **1** (1988) 255–299.
9. Faltings, G., *Crystalline cohomology and p-adic Galois representations*, in Algebraic analysis, geometry and number theory. Proceedings of the JAMI Inaugural Conference, J. I. Igusa, ed., Johns Hopkins University Press, Baltimore (1989) 25–80.
10. Flach, M., *A finiteness theorem for the symmetric square of an elliptic curve*, Invent. math. **109** (1992) 307–327.
11. Frey, G., *Links between solutions of $A - B = C$ and elliptic curves*, in Number theory, Ulm 1987, Proceedings, Lecture Notes in Math. **1380**, Springer-Verlag, New York (1989) 31–62.
12. Gelbart, S., Automorphic forms on adele groups, Ann. of Math. Studies **83**, Princeton University Press, Princeton (1975).
13. Gross, B., *Kolyvagin's work on modular elliptic curves*, in L-functions and arithmetic, London Math. Soc. Lecture Notes **153** (1991) 235–256.
14. Hardy, G. H., Wright, E. M., An introduction to the theory of numbers, Fourth edition, Oxford University Press, London (1971).
15. Hartshorne, R., Algebraic geometry, Springer-Verlag, New York (1977).
16. Hellegouarch, Y., *Étude des points d'ordre fini des variétés de dimension un définies sur un anneau principal*, J. Reine Angew. Math. **244** (1970) 20–36.
17. Hellegouarch, Y., *Points d'ordre fini des variétés de dimension un*, in Colloque de Théorie des Nombres (Univ. Bordeaux, Bordeaux, 1969), Bull. Soc. Math. France, Mém. **25**, Soc. Math. France, Paris (1971) 107–112.
18. Hellegouarch, Y., *Points d'ordre fini sur les courbes elliptiques*, C. R. Acad. Sci. Paris Sér. A-B **273** (1971) A540–A543.

19. Hellegouarch, Y., *Points d'ordre* $2p^h$ *sur les courbes elliptiques*, Acta. Arith. **26** (1974/75) 253–263.
20. Kolyvagin, V. A., *Euler systems*, in The Grothendieck Festschrift (Vol. II), P. Cartier, et. al., eds., Birkhäuser, Boston (1990) 435–483.
21. Langlands, R., Base change for GL(2), Ann. of Math. Studies **96**, Princeton University Press, Princeton (1980).
22. Mazur, B., *Deforming Galois representations*, in Galois groups over $\mathbf{Q}$, Y. Ihara, K. Ribet, J-P. Serre, eds., MSRI publications **16**, Springer-Verlag, New York (1989) 385–437.
23. Mazur, B., *Number theory as gadfly*, Amer. Math. Monthly **98** (1991) 593–610.
24. Mazur, B., Tilouine, J., *Représentations galoisiennes, différentielles de Kähler et "conjectures principales"*, Publ. Math. IHES **71** (1990) 65–103.
25. Oesterlé, J., *Nouvelles approches du "théorème" de Fermat*, Séminaire Bourbaki n^o 694 (1987–88), Astérisque **161–162** (1988) 165–186.
26. Ramakrishna, R., *On a variation of Mazur's deformation functor*, Compositio Math. **87** (1993) 269-286.
27. Ribenboim, P., 13 lectures on Fermat's Last Theorem, Springer-Verlag, New York (1979).
28. Ribet, K., *On modular representations of* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. math. **100** (1990) 431–476.
29. Ribet, K. *Report on mod* ℓ *representations of* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, preprint.
30. Rubin, K., The main conjecture. Appendix to: Cyclotomic fields I and II, S. Lang, Graduate Texts in Math. **121**, Springer-Verlag, New York (1990) 397–419.
31. Rubin, K., *Kolyvagin's system of Gauss sums*, in Arithmetic algebraic geometry, G. van der Geer, F. Oort, and J. Steenbrink, eds., Progress in Math. **89**, Birkhäuser, Boston (1991) 309–324.
32. Rubin, K., *The "main conjectures" of Iwasawa theory for imaginary quadratic fields*, Invent. math. **103** (1991) 25–68.
33. Serre, J-P., *Sur les représentations modulaires de degré 2 de* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. **54** (1987) 179–230.
34. Shimura, G., *Correspondances modulaires et les fonctions* ζ *de courbes algébriques*, J. Math. Soc. Japan **10** (1958) 1–28.
35. Shimura, G., *Construction of class fields and zeta functions of algebraic curves*, Ann. of Math. **85** (1967) 58–159.
36. Shimura, G., Introduction to the arithmetic theory of automorphic functions, Princeton Univ. Press, Princeton (1971).
37. Shimura, G., *On elliptic curves with complex multiplication as factors of the Jacobians of modular function fields*, Nagoya Math. J. **43** (1971) 199–208.
38. Shimura, G., *On the factors of the jacobian variety of a modular function field*, J. Math. Soc. Japan **25** (1973) 523–544.
39. Shimura, G., *Yutaka Taniyama and his time. Very personal recollections*, Bull. London Math. Soc. **21** (1989) 186–196.
40. Silverman, J., The arithmetic of elliptic curves, Graduate Texts in Math. **106**, Springer-Verlag, New York (1986).
41. Thaine, F., *On the ideal class groups of real abelian number fields*, Ann. of Math. **128** (1988) 1–18.
42. Tunnell, J., *Artin's conjecture for representations of octahedral type*, Bull. A.M.S. **5** (1981) 173–175.
43. Weil, A., *Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen*, Math. Ann. **168** (1967) 149–156.

Department of Mathematics, Ohio State University, Columbus, Ohio 43210
E-mail address: rubin@math.ohio-state.edu

Department of Mathematics, Ohio State University, Columbus, Ohio 43210
E-mail address: silver@math.ohio-state.edu