BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 53, Number 1, January 2016, Pages 1–39
http://dx.doi.org/10.1090/bull/1515
Article electronically published on August 25, 2015

RECIPROCITY LAWS AND GALOIS REPRESENTATIONS: RECENT BREAKTHROUGHS

JARED WEINSTEIN

Abstract. Given a polynomial $f(x)$ with integer coefficients, a *reciprocity law* is a rule which determines, for a prime p, whether $f(x)$ modulo p is the product of distinct linear factors. We examine reciprocity laws through the ages, beginning with Fermat, Euler and Gauss, and continuing through the modern theory of modular forms and Galois representations. We conclude with an exposition of Peter Scholze's astonishing work on torsion classes in the cohomology of arithmetic manifolds.

1. Introduction

1.1. **Motivation: the splitting problem.** Suppose $f(x)$ is a monic irreducible polynomial with integer coefficients. If p is a prime number, then reducing the coefficients of $f(x)$ modulo p gives a new polynomial $f_p(x)$, which may be reducible. We say that $f(x)$ is *split modulo* p if $f_p(x)$ is the product of distinct linear factors.

Example 1.1.1. The polynomial $f(x) = x^2 + 1$ is split modulo 5, since $f_5(x) \equiv (x+2)(x+3)$ mod 5. But it is not split modulo 7, since $f_7(x)$ is irreducible, nor is it split modulo 2, since $f_2(x) \equiv (x+1)^2 \pmod 2$ has a repeated factor. The first few p for which $x^2 + 1$ is split modulo p are $5, 13, 17, 29, 37, 41, 53, \ldots$.

Example 1.1.2. The polynomial $f(x) = x^3 - 2$ is split modulo 31, since $f_{31}(x) \equiv (x+11)(x+24)(x+27) \pmod{31}$. But it is not split modulo 5, since $f_5(x) \equiv (x+2)(x^2+3x+4) \pmod 5$, and the second factor is irreducible. The first few p for which $x^3 - 2$ is split modulo p are $31, 43, 109, 127, 157, 223, 229, \ldots$.

This article is concerned with the following simple question.

Question A. Given an irreducible polynomial $f(x)$ with integer coefficients, is there a rule which, for every prime p, determines whether $f(x)$ is split modulo p?

A large swath of modern number theory known as the *Langlands program* is dedicated to variations on the theme of Question A.

We ought to clarify what is meant by a "rule" in Question A. We are not looking for an algorithm to factor a polynomial modulo a prime. Rather we are seeking a systematic connection to some other part of mathematics. Such a rule is called a *reciprocity law*. Our search for reciprocity laws can be rephrased as the study of a single group, the absolute Galois group of the field of rational numbers, written $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. The representation theory of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ has been particularly fruitful in answering instances of Question A. In this article we will review reciprocity laws in four successive epochs:

Received by the editors May 18, 2015.
2010 *Mathematics Subject Classification.* Primary 11R37, 11R39, 11F80.

1. The solution of Question A in the case of $f(x) = x^2 + 1$ is due to Fermat. The solution for a general quadratic polynomial was conjectured by Euler and first proved by Gauss; this is the famous quadratic reciprocity law.
2. Thereafter, many other reciprocity laws followed, due to Eisenstein, Kummer, Hilbert, Artin, and others, leading up to the formulation of *class field theory* in the early 20th century. These reciprocity laws are called *abelian*. They only apply to those instances of Question A where the polynomial $f(x)$ has a solvable Galois group.
3. In the second half of the 20th century, a remarkable link was found between *modular forms* and two-dimensional representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, due to Eichler, Shimura, Deligne, and Serre. This made it possible to find reciprocity laws for certain quintic $f(x)$ with nonsolvable Galois group.
4. The 21st century has seen an explosion of results which link representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ to the geometry of *arithmetic manifolds*. We highlight Scholze's recent work [Sch13c], which employs techniques invented within the past five years.

Besides the exposition of [Sch13c] there is much overlap between this article and other surveys about reciprocity laws. Our Question A is lifted almost verbatim from B. Wyman's 1972 article [Wym72], which contains a brief introduction to algebraic number theory. The article [AG00] is an exposition of reciprocity laws in the context of Fermat's Last Theorem. C. Dalawat's essay [SRY12, Ch. 2] describes the link between reciprocity laws and modular forms, with many examples.

2. Fermat, Euler, and Gauss

2.1. **Quadratic reciprocity laws.** Which positive integers n are the sum of two squares? Fermat settled this question in 1640. Using his method of "descent", he showed that if a prime number p divides a sum of two squares, neither of which is divisible by p, then p is itself a sum of two squares. Also one sees from the identity $(a^2+b^2)(c^2+d^2) = (ad-bc)^2 + (ac+bd)^2$ that the property of being a sum of two squares is preserved under multiplication. From there it is simple to check that n is a sum of two squares if and only if $n = p_1 \cdots p_k m^2$, where each of the primes $p_1, \ldots, p_k$ is a sum of two squares, and $m \geq 1$.

Thus we are reduced to the case that $n = p$ is prime. We already mentioned that p is a sum of two squares if it divides a sum of two squares, neither of which is divisible by p. Thus we are trying to determine when the congruence $a^2 + b^2 \equiv 0 \pmod{p}$ has a solution for $a, b \not\equiv 0 \pmod{p}$. Recall that the ring $\mathbf{Z}/p\mathbf{Z}$ of integers modulo p is a field. After dividing by b^2 and relabeling, this becomes $x^2 + 1 \equiv 0 \pmod{p}$. Deciding when it can be solved turns out to be equivalent to answering Question A for $f(x) = x^2 + 1$.

Theorem 2.1.1. *Let p be an odd prime. Then $x^2 + 1 \equiv 0 \pmod{p}$ has a solution if and only if $p \equiv 1 \pmod{4}$.*

Proof. Suppose $x^2 + 1 \equiv 0 \pmod{p}$. Then $x^{p-1} = (x^2)^{(p-1)/2} \equiv (-1)^{(p-1)/2} \pmod{p}$. But by Fermat's Little Theorem, $x^{p-1} \equiv 1 \pmod{p}$, implying that $(-1)^{(p-1)/2} = 1$ and therefore $p \equiv 1 \pmod{4}$.

Conversely, suppose $p \equiv 1 \pmod{4}$. Let $x = ((p-1)/2)!$. We have $x^2 \equiv (-1)^{(p-1)/2}(p-1)! \pmod{p}$ (by pairing up n with $-n$ in the product), which is $(p-1)! \pmod{p}$, and by Wilson's theorem this is $\equiv -1 \pmod{p}$. □

Another way of phrasing Theorem 2.1.1 is that x^2+1 splits modulo a prime p if and only if $p \equiv 1 \pmod 4$. (Note that modulo 2, $x^2+1 \equiv (x+1)^2$ contains a repeated root, and so is not split as we have defined it. Given an irreducible polynomial $f(x)$, the primes p for which $f_p(x)$ has a repeated factor all divide the discriminant of $f(x)$, and hence are finite in number.)

Theorem 2.1.1 demonstrates the simplest possible sort of reciprocity law, namely one where the factorization of $f(x)$ modulo p is determined by a *congruence condition on p*. As the following examples show, this is also the case for other quadratic polynomials.

Example 2.1.2. The polynomial $f(x) = x^2+x+1$ splits modulo p if and only if $p \equiv 1 \pmod 3$. Let us sketch a proof of this fact. In one direction: If $p \equiv 1 \pmod 3$, then 3 divides $p-1$, which is the order of the group $(\mathbf{Z}/p\mathbf{Z})^\times$. By Cauchy's theorem there exists an element $w \in (\mathbf{Z}/p\mathbf{Z})^\times$ of order 3. Then w is a root of $x^2+x+1 = (x^3-1)/(x-1)$. In the other direction: We can rule out $p=3$ since $f(x) \equiv (x-1)^2 \pmod 3$. If x^2+x+1 has a root $w \pmod p$, then $w^3 \equiv 1$ but $w \not\equiv 1 \pmod p$, so that $(\mathbf{Z}/p\mathbf{Z})^\times$ contains an element of order 3, and thus 3 divides $p-1$.

Example 2.1.3. The polynomial $f(x) = x^2-2$ splits modulo p if and only if $p \equiv \pm 1 \pmod 8$. We will only prove part of this fact: Suppose that $p \equiv 1 \pmod 8$. We now apply a theorem from elementary number theory which tells us that $(\mathbf{Z}/p\mathbf{Z})^\times$ is a *cyclic group* of order $p-1$. Let g be a generator of $(\mathbf{Z}/p\mathbf{Z})^\times$, and let $y = g^{(p-1)/8}$, $x = y+y^{-1}$. Then $y^4 = -1$ and therefore $y^2 = -y^{-2}$; thus $x^2 = y^2+2+y^{-2} = 2$. This proof is based on the identity of complex numbers $e^{2\pi i/8}+e^{-2\pi i/8} = \sqrt{2}$. Note that $e^{2\pi i/8}$ is a primitive 8th root of 1; its analogue in $\mathbf{Z}/p\mathbf{Z}$ is what we have called y.

Example 2.1.4. Is there a similar rule for the polynomial $f(x) = x^2-5$? Note that if $f_p(x)$ factors into linear factors, then there is an integer n such that $n \pmod p$ is a root of $f_p(x)$, so p divides n^2-5. Conversely, if p divides n^2-5, then $f_p(x) \equiv (x-n)(x+n) \pmod p$. Thus in the table below $f_p(x)$ splits for each red prime p.

n	Factorization of $f(n)$	n	Factorization of $f(n)$
1	-2^2	8	61
2	-1	9	$2^2 \cdot 19$
3	2^2	10	$5 \cdot 19$
4	11	11	$2^2 \cdot 29$
5	$2^2 5$	12	139
6	31	13	$2^2 \cdot 41$
7	$2^2 \cdot 11$	14	191

(We are ignoring 2 and 5, since these are the prime divisors of the discriminant of $f(x)$.) The red primes are all congruent to 1 modulo 5.

In fact the p for which $f_p(x)$ splits are described by a congruence condition whenever $f(x)$ is a quadratic polynomial:

Theorem 2.1.5 (Quadratic reciprocity)**.** *Let $f(x) = x^2+bx+c$ be a monic irreducible polynomial with integer coefficients, so $d = b^2-4c$ is not a square. Then for p not dividing d, the splitting behavior of $f(x)$ modulo p is determined by the congruence class of p modulo d.*

This is a form of the *quadratic reciprocity law*, which was conjectured by Euler and proved (many times over) by Gauss. If p is an odd prime, then $f(x)$ factors modulo p if and only if d is congruent to a square modulo p. We define the *Legendre symbol* $(\frac{d}{p})$ for any odd prime p and any integer d prime to p as 1 if d is a square modulo p and -1 otherwise. Thus for instance Theorem 2.1.1 is the statement that $(\frac{-1}{p}) = (-1)^{(p-1)/2}$. In elementary number theory texts there appears a more precise version of Theorem 2.1.5: If $q \neq p$ is an odd prime, then

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}},$$

which implies that the splitting behavior of $x^2 - q$ modulo p depends on the congruence class of p modulo $4q$. The symmetry between p and q explains the term "reciprocity" for such laws.

Let us return for a moment to Fermat's theorem on sums of squares. Could it apply to the representation of integers by other quadratic forms, such as $a^2 + 5b^2$? Theorem 2.1.5 shows that a prime $p \neq 2, 5$ divides an integer of the form $x^2 + 5$ if and only if p satisfies a congruence condition modulo 20, which happens to be the condition that $p \equiv 1, 3, 7, 9 \pmod{20}$. But such a prime (for instance 7) is not necessarily of the form $a^2 + 5b^2$. It turns out that Fermat's method of descent fails in this context. Phrased in modern terms, the culprit is the failure of $\mathbf{Z}[\sqrt{-5}]$ to have the property of unique factorization into primes. In fact $p = a^2 + 5b^2$ if and only if $p \equiv 1, 9 \pmod{20}$. For a fascinating account of the problem of classifying primes of the form $x^2 + ny^2$, see Cox's book of the same title [Cox89].

2.2. **Some reciprocity laws of higher degree.** What about polynomials $f(x)$ of higher degree? A little experimentation will reveal that the factorization behavior of a "random" cubic or quartic polynomial will be influenced, but not completely determined, by a congruence condition modulo p. For instance, in Example 1.1.2, the primes for which $x^3 - 2$ is split are all congruent to 1 modulo 3, but the converse is false. There are special cases where a congruence condition is the complete story: for instance the polynomial $x^3 + x^2 - 2x - 1$ splits modulo p if and only if $p \equiv \pm 1 \pmod{7}$. When is the splitting behavior of a polynomial determined by congruence conditions?

For a clue, let $m \geq 1$, and consider the polynomial $x^m - 1$. If p does not divide m, $x^m - 1$ splits modulo p if and only if the multiplicative group $(\mathbf{Z}/p\mathbf{Z})^\times$ contains m distinct elements of order dividing m. Since $(\mathbf{Z}/p\mathbf{Z})^\times$ is a cyclic group of order $p - 1$, this happens exactly when $p \equiv 1 \pmod{m}$. This logic extends to show that the splitting behavior of $f(x)$ is determined by congruence conditions whenever $f(x)$ is a factor of $x^m - 1$. Using some algebraic number theory, it can be shown that splitting is based on congruence conditions modulo m for those $f(x)$ whose roots are contained in the *cyclotomic field* $\mathbf{Q}(\zeta_m)$, where $\zeta_m = \exp(2\pi i/m)$. For instance, the roots of $x^3 + x^2 - 2x - 1$ are $\zeta_7^k + \zeta_7^{-k}$, where $k = 1, 2, 3$, which explains why the splitting behavior of this polynomial modulo p is determined by p modulo 7.

Thus there is a satisfactory answer to Question A whenever the roots of $f(x)$ are contained in a cyclotomic field. Surprisingly, the converse is also true. See [Wym72] for a discussion of the proof of the following theorem.

Theorem 2.2.1. *The splitting behavior of $f(x)$ modulo p is determined by congruence conditions on p if and only if the roots of $f(x)$ are contained in a cyclotomic field.*

What would a reciprocity law look like if it isn't a set of congruence conditions?

Example 2.2.2. Let $f(x) = x^3 - 2$. If $f(x)$ splits modulo p, then 2 has distinct cube roots x_1, x_2, x_3 modulo p. The ratio x_1/x_2 must have order 3 in $(\mathbf{Z}/p\mathbf{Z})^\times$. So 3 divides $|(\mathbf{Z}/p\mathbf{Z})^\times| = p-1$, and thus $p \equiv 1 \pmod 3$. Here are the first few primes $p \equiv 1 \pmod 3$, with the primes for which $f(x)$ splits shown in red:

7, 13, 19, 31, 37, 43, 61, 67, 73, 79, 97, 103, 109, 127, 139, 151, 157, 163, 181, 193, 199, 211, 223, 229, 241, ...

There does not seem to be a pattern arising from congruence conditions on p. In fact Theorem 2.2.1 shows that there cannot be one, since $\mathbf{Q}(\sqrt[3]{2})$ is not contained in a cyclotomic field.

A similar analysis can be carried out for the polynomial $f(x) = x^4 - 2$. In order for this polynomial to split modulo p, it is necessary but not sufficient that $p \equiv 1 \pmod 8$. Reciprocity laws for both $x^3 - 2$ and $x^4 - 2$ were conjectured by Euler and proved by Gauss.

Theorem 2.2.3 ([Cox89, Theorems 4.15 and 4.23(ii)])**.** *The polynomial $x^3 - 2$ splits modulo p if and only if $p = a^2 + 27b^2$ for integers a and b. The polynomial $x^4 - 2$ splits modulo p if and only if $p = a^2 + 64b^2$ for integers a and b.*

Unlike the case of $a^2 + b^2$, which represents p if and only if $p \equiv 1 \pmod 4$, the representation of p by the quadratic forms $a^2 + 27b^2$ and $a^2 + 64b^2$ is not determined by a congruence condition on p. But in fact there are disguised congruence conditions in Theorem 2.2.3, which were well known to Gauss. Let us focus on $x^4 - 2$. If $x^4 - 2$ splits modulo p, then the quotient of two of its roots in $\mathbf{Z}/p\mathbf{Z}$ must be a square root of -1, so that by Theorem 2.1.1 we have $p \equiv 1 \pmod 4$. By Fermat's theorem $p = a^2 + b^2$. Without loss of generality, assume that a is odd and b is even. We now pass to the ring $\mathbf{Z}[i]$ of *Gaussian integers*, the subring of $\mathbf{C}$ consisting of those $a + bi$ with $a, b \in \mathbf{Z}$. In $\mathbf{Z}[i]$, p is no longer prime; we have $p = \pi\overline{\pi}$, where $\pi = a + bi$. Theorem 2.2.3 says that $x^4 - 2$ splits modulo p if and only if $\pi \equiv 1, 3, 5, 7 \pmod{8\mathbf{Z}[i]}$. Indeed, this condition translates into the statement that $b = 8b_0$ for an integer b_0, so $p = a^2 + 64b_0^2$. Thus the splitting behavior of $x^4 - 2$ modulo a prime $p \equiv 1 \pmod 4$ is determined by a congruence condition on a prime of $\mathbf{Z}[i]$ which divides p.

As an example, $13 = (3 + 2i)(3 - 2i)$. But $3 + 2i$ is not congruent to $1, 3, 5,$ or $7 \pmod{8\mathbf{Z}[i]}$, and therefore $x^4 - 2$ is not split modulo 13. On the other hand $73 = (3 + 8i)(3 - 8i)$, and $3 + 8i \equiv 3 \pmod{8\mathbf{Z}[i]}$, so that $x^4 - 2$ is split modulo 73.

The analysis for $x^3 - 2$ is similar, but involves the *Eisenstein integers* $\mathbf{Z}[\omega]$, where $\omega = e^{2\pi i/3}$. For a discussion of reciprocity laws for polynomials of the form $x^3 - a$ and $x^4 - a$, including a proof of Theorem 2.2.3, see [IR90, Ch. 9].

3. Class field theory

3.1. **Some algebraic number theory.** At this point it is appropriate to introduce some basic notions from algebraic number theory. If $f(x)$ is an irreducible

polynomial with rational coefficients, then $K = \mathbf{Q}[x]/f(x)$ is an *algebraic number field*. Let $\mathcal{O}_K$ be the integral closure of $\mathbf{Z}$ in K. It is a basic fact of algebraic number theory that $\mathcal{O}_K$ is a *Dedekind domain*. This means that even though $\mathcal{O}_K$ may not have the property of unique factorization, it does have the corresponding property for ideals. See for instance [IR90, Ch. 12]. It is common to refer to a nonzero prime ideal of $\mathcal{O}_K$ as a "prime of K" or a "finite place of K".

Example 3.1.1. Let $K = \mathbf{Q}(\sqrt{-5})$. Then $\mathcal{O}_K = \mathbf{Z}[\sqrt{-5}]$. The element 6 admits two factorizations $2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$ into irreducible elements of $\mathcal{O}_K$, none of which divide any other. However, every nonzero ideal in $\mathcal{O}_K$ factors into prime ideals in one way only. The ideals $P = (2, 1+\sqrt{-5})$, $Q = (3, 1+\sqrt{-5})$ and $\overline{Q} = (3, 1-\sqrt{-5})$ of $\mathbf{Z}[\sqrt{-5}]$ are all prime, and we have $(2) = P^2$, $(3) = Q\overline{Q}$, $(1+\sqrt{-5}) = PQ$ and $(1-\sqrt{-5}) = P\overline{Q}$. The ideal (6) has a unique factorization into prime ideals, namely $P^2 Q\overline{Q}$.

If p is a prime number, then $p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_k^{e_k}$ is a product of powers of distinct primes of K. For each i, the ring $\mathcal{O}_K/\mathfrak{p}_i$ is a finite field extension of $\mathbf{Z}/p\mathbf{Z}$. If $e_i = 1$ for all i, we call p *unramified* in K. We say that p is *split* in K if it is unramified and if $\mathcal{O}_K/\mathfrak{p}_i \cong \mathbf{Z}/p\mathbf{Z}$ for each i. If $e_i > 1$ for some i, then p is *ramified* in K. There are finitely many ramified primes.

The splitting of primes in number fields is closely related to the splitting of polynomials modulo primes. Let $f(x)$ be a monic irreducible polynomial with integer coefficients, and let $K = \mathbf{Q}(\alpha)$ be the field obtained by adjoining to $\mathbf{Q}$ a single root α of $f(x)$. Then with possibly finitely many exceptions, $f(x)$ is split modulo p if and only if p splits in K (see [Lan94, Proposition 27]).

These notions have analogues in an extension of number fields L/K. If $\mathfrak{p}$ is a prime of K, then we can factor $\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_k^{e_k}$ into powers of distinct primes of L. The primes $\mathfrak{P}_i$ are said to divide $\mathfrak{p}$. We say $\mathfrak{p}$ is *unramified* in L if each $e_i = 1$, and *ramified* otherwise; there are finitely many ramified primes. We say $\mathfrak{p}$ is *split*[1] in L if it is unramified and if for each i, $\mathcal{O}_L/\mathfrak{P}_i \cong \mathcal{O}_K/\mathfrak{p}$.

The "correct" generalization of Question A is then:

Question B. Let L/K be an extension of number fields. Is there a rule for determining when a prime ideal of K is split in L?

Question B is inextricably linked with Galois theory. Recall that if K is a field, an extension L/K is *Galois* if it is the splitting field of a collection of separable polynomials with coefficients in K (separable means no repeated roots). If L/K is Galois, the Galois group $\mathrm{Gal}(L/K)$ is the group of field automorphisms of L which act as the identity on K. Its cardinality is the same as the degree of L/K (that is, the dimension of K as an L-vector space). The philosophy of Galois theory is that there is no algebraic means of distinguishing the roots of an irreducible polynomial within the field they generate (such as $\sqrt{2}$ and $-\sqrt{2}$ within $\mathbf{Q}(\sqrt{2})$), and that one can bring to bear the power of group theory in analyzing those roots.

Example 3.1.2. Let K be a finite field of cardinality q, and let L/K is a finite extension of degree d. Then L/K is Galois. The group $\mathrm{Gal}(L/K)$ is a cyclic group of order d, with generator $x \mapsto x^q$.

[1] What we have defined as split, other authors sometimes call *completely split*.

Example 3.1.3. For an integer $m \geq 1$, we have the cyclotomic field $\mathbf{Q}(\zeta_m)$. There is an isomorphism $\alpha\colon \mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q}) \to (\mathbf{Z}/m\mathbf{Z})^\times$ characterized by $\sigma(\zeta_m) = \zeta_m^{\alpha(\sigma)}$ for each $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q})$. Note that if c is complex conjugation, then $c(\zeta_m) = \overline{\zeta}_m = \zeta_m^{-1}$, so that $\alpha(c) \equiv -1 \pmod{m}$.

Example 3.1.4. The splitting field of the polynomial x^4-2 over $\mathbf{Q}$ is $L = \mathbf{Q}(i, \sqrt[4]{2})$. The Galois group $\mathrm{Gal}(L/\mathbf{Q})$ is the dihedral group of order 8, generated by two elements r and s, defined by the table

$$\begin{aligned} r(\sqrt[4]{2}) &= i\sqrt[4]{2}, & s(\sqrt[4]{2}) &= \sqrt[4]{2}, \\ r(i) &= i, & s(i) &= -i. \end{aligned}$$

These generators satisfy the relations $r^4 = 1$, $s^2 = 1$, and $srs^{-1} = r^{-1}$.

Let us now return to Question B. If L/K is an arbitrary extension of number fields, let L'/K be a *Galois closure*. This is a Galois extension of minimal degree containing L. It turns out that a prime of K is split in L if and only if it is split in L' (see [Neu99, §8, Exercise 4]). Thus to answer Question B, it suffices to assume that L/K is Galois.

Let L/K be Galois extension of number fields, let $\mathfrak{p}$ be a prime of K, and let $\mathfrak{P}$ be a prime of L dividing $\mathfrak{p}$. The number of elements of $\mathcal{O}_K/\mathfrak{p}$ is denoted $N\mathfrak{p}$. If $\mathfrak{p}$ is unramified in L, then there exists a distinguished automorphism $\mathrm{Frob}_{\mathfrak{P}|\mathfrak{p}} \in \mathrm{Gal}(L/K)$ called the *Frobenius automorphism* which is characterized by the relation

$$\mathrm{Frob}_{\mathfrak{P}|\mathfrak{p}}(x) \equiv x^{N\mathfrak{p}} \bmod \mathfrak{P}$$

for all $x \in \mathcal{O}_L$.

If $\mathfrak{P}'$ is another prime of L dividing $\mathfrak{p}$, the automorphisms $\mathrm{Frob}_{\mathfrak{P}|\mathfrak{p}}$ and $\mathrm{Frob}_{\mathfrak{P}'|\mathfrak{p}}$ are conjugate in $\mathrm{Gal}(L/K)$. Thus one can talk about a well-defined *conjugacy class* $\mathrm{Frob}_\mathfrak{p}$ in $\mathrm{Gal}(L/K)$. If L/K happens to be abelian, then $\mathrm{Frob}_\mathfrak{p}$ is a well-defined element of $\mathrm{Gal}(L/K)$. An important observation is that for a prime $\mathfrak{p}$ of K that is unramified in L,

$$\mathrm{Frob}_\mathfrak{p} = 1 \text{ in } \mathrm{Gal}(L/K) \text{ if and only if } \mathfrak{p} \text{ is split in } L.$$

This criterion makes sense even when L/K is not abelian, since the conjugacy class of the identity always has one element.

Example 3.1.5. Let $K = \mathbf{Q}$, $L = \mathbf{Q}(i)$. Then $\mathrm{Gal}(L/K) = \{1, c\}$, where c is complex conjugation. The prime 3 remains prime in $\mathbf{Z}[i]$. Then $\mathrm{Frob}_3 = c$, since for $a, b \in \mathbf{Z}$,

$$(a+bi)^3 \equiv a^3 + b^3 i^3 \equiv a - bi \equiv \overline{a+bi} \pmod{3\mathbf{Z}[i]}.$$

On the other hand, 5 splits in $\mathbf{Q}(i)$ as $(5) = (2+i)(2-i)$. Since

$$(a+bi)^5 \equiv a^5 + b^5 i^5 \equiv a + bi \pmod{2+i},$$

we can conclude $\mathrm{Frob}_5 = 1$.

3.2. **The reciprocity map.** *Class field theory* refers to the complete solution of Question B in the case that L/K is Galois and $\mathrm{Gal}(L/K)$ is *abelian*. Such extensions are simply called *abelian*. Roughly speaking, it predicts that for a prime $\mathfrak{p}$ of K which is unramified in L, the element $\mathrm{Frob}_\mathfrak{p} \in \mathrm{Gal}(L/K)$ is determined by "congruence conditions" on $\mathfrak{p}$.

To make a precise statement we need a few definitions. Recall that a nonzero prime ideal of $\mathcal{O}_K$ is also called a finite place of K. An *infinite place* of K is a field

embedding $\tau\colon K \hookrightarrow \mathbf{C}$, considered up to complex conjugation. An infinite place is either a real place or a complex place, as $\tau(K)$ is contained in $\mathbf{R}$ or not.

Theorem 3.2.1. *Let L/K be an abelian extension of number fields. There exists an ideal $\mathfrak{f}$ of $\mathcal{O}_K$ (depending on L) with the following property. Suppose $\mathfrak{p} = (\pi)$ is a principal prime ideal of K such that $\pi \equiv 1 \pmod{\mathfrak{f}}$, and such that $\tau(\pi) > 0$ for all real places τ of K. Then $\mathfrak{p}$ is split in L.*

We remark that there is a refined version of Theorem 3.2.1 involving a homomorphism called the *reciprocity map* or *Artin map*, which generalizes the Legendre symbol. See [Cas67, Ch. VII].

Example 3.2.2. In the cases $K = \mathbf{Q}$ and $L = \mathbf{Q}(i)$, we can verify that Theorem 3.2.1 holds with $\mathfrak{f} = (4)$: this reduces to the statement that if $p \equiv 1 \pmod 4$ is a (positive) prime, then p splits in $\mathbf{Q}(i)$.

Example 3.2.3. More generally, suppose $K = \mathbf{Q}$ and $L/\mathbf{Q}$ is an abelian extension. In the context of Theorem 3.2.1, write $\mathfrak{f} = (m)$, where $m \geq 1$. Then any $p \equiv 1 \pmod m$ splits in L. Compare this with the behavior of the cyclotomic field $\mathbf{Q}(\zeta_m)$: a prime p splits in $\mathbf{Q}(\zeta_m)$ if and only if $p \equiv 1 \pmod m$. Thus the set of primes which split completely in $\mathbf{Q}(\zeta_m)$ is contained in the set of primes which split completely L. This fact can be used to show that L is contained in $\mathbf{Q}(\zeta_m)$. See for instance [Cas67, Theorem 6.1]. Therefore we have

Theorem 3.2.4 (The Kronecker–Weber theorem, [Was97, Thm. 14.1])**.** *Every extension of $\mathbf{Q}$ with an abelian Galois group is contained in a cyclotomic field.*

Example 3.2.5. Let $K = \mathbf{Q}(i)$ and $L = K(\sqrt[4]{2})$. It turns out that in Theorem 3.2.1 we can take $\mathfrak{f} = (8)$. Suppose p is a prime number of the form $a^2 + 64b^2$, where $a \equiv 1 \pmod 8$. Then $p = \pi\overline{\pi}$ where $\pi = a + bi$ is a prime in $\mathbf{Z}[i]$ and $\pi \equiv 1 \pmod{8\mathbf{Z}[i]}$, and we can deduce that p splits in L. This recovers part of Theorem 2.2.3.

Theorem 3.2.1 is the work of many people, including Artin, Hasse, Furtwängler, Takagi, and others. It allows us to answer Question B in the case that the polynomial $f(x)$ is *solvable*, meaning that its roots lie in a tower of number fields $\mathbf{Q} = K_0 \subset K_1 \subset \cdots \subset K_n = K$, with each K_{i+1}/K_i abelian. A prime p splits in K if and only if $\mathfrak{p}$ splits in K_1, a prime dividing $\mathfrak{p}$ in K_1 splits in K_2, and so on, with each splitting being governed by congruences. In Example 3.2.5, the relevant tower was $\mathbf{Q} \subset \mathbf{Q}(i) \subset \mathbf{Q}(i, \sqrt[4]{2})$.

Not all extensions of number fields are solvable. For instance, if $f(x)$ is a "random" quintic polynomial with rational coefficients, then the Galois group of f is likely to be S_5, which contains the nonabelian simple group A_5. Theorem 3.2.1 makes no predictions about the splitting behavior of primes in an A_5-extension. The first "nonsolvable reciprocity laws" were discovered by Shimura in the 1960s [Shi66], and further investigated by Deligne [Del71] and Deligne and Serre [DS74]. These reciprocity laws link Galois representations with modular forms.

3.3. **The absolute Galois group of a number field, and Galois representations.** It is immensely useful to talk about all of the finite extensions of a number field K at once, as living in an algebraic closure $\overline{K}$. This leads to the *absolute Galois group* $\mathrm{Gal}(\overline{K}/K)$, which is the group of automorphisms of $\overline{K}$ which act as

the identity on K. We have

$$\mathrm{Gal}(\overline{K}/K) = \varprojlim_{L} \mathrm{Gal}(L/K),$$

where L runs over finite Galois extensions of K. Written this way, $\mathrm{Gal}(\overline{K}/K)$ becomes a topological group, whose open subgroups are exactly the subgroups $\mathrm{Gal}(\overline{K}/L)$ consisting of automorphisms which act trivially on a finite extension L/K. Focus can then shift from particular number fields L/K to the topological group $\mathrm{Gal}(\overline{K}/K)$.

The group $\mathrm{Gal}(\overline{K}/K)$ is very complicated. It is difficult even to write down particular elements of it (Zorn's lemma is usually required). The best way to study $\mathrm{Gal}(\overline{K}/K)$ is through *Galois representations*, which for our purposes are continuous homomorphisms

$$\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_n(F),$$

where F is some topological field. We say that ρ is *unramified* at a prime $\mathfrak{p}$ of K if it factors through $\mathrm{Gal}(L/K)$, where L/K is some (possibly infinite) algebraic extension which is unramified at $\mathfrak{p}$. In that case $\rho(\mathrm{Frob}_{\mathfrak{p}})$ is a well-defined conjugacy class in $\mathrm{GL}_n(F)$.

Question C. Given a Galois representation $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_n(F)$, is there a rule for determining the conjugacy class of $\rho(\mathrm{Frob}_{\mathfrak{p}})$ for the unramified primes $\mathfrak{p}$?[2]

3.4. **Artin representations.** If $E = \mathbf{C}$ is the field of complex numbers, then a Galois representation $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_n(\mathbf{C})$ is called an *Artin representation*. The image of an Artin representation is necessarily finite, so that ρ factors through a representation of a finite group $\mathrm{Gal}(L/K)$. Solving Question C for Artin representations would also solve Question B. For instance if L/K is finite and Galois, then one can find an Artin representation ρ whose kernel is exactly $\mathrm{Gal}(\overline{K}/L)$, and then a prime $\mathfrak{p}$ splits in L if and only if $\rho(\mathrm{Frob}_{\mathfrak{p}}) = 1$.

Example 3.4.1 (One-dimensional Artin representations over $\mathbf{Q}$)**.** A one-dimensional Artin representation over the base field $\mathbf{Q}$ is just a continuous homomorphism $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{C}^\times$. By the Kronecker–Weber theorem (Theorem 3.2.4), there is an m for which ρ factors through $\mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q}) \cong (\mathbf{Z}/m\mathbf{Z})^\times$. Thus one-dimensional Artin representations ρ correspond to *Dirichlet characters*, which are complex-valued characters χ of $(\mathbf{Z}/m\mathbf{Z})^\times$. This correspondence is characterized by the relation $\rho(\mathrm{Frob}_p) = \chi(p)$ for all p not dividing m.

Example 3.4.2 (A dihedral Artin representation)**.** Recall from Example 3.1.4 that $\mathrm{Gal}(\mathbf{Q}(i, \sqrt[4]{2})/\mathbf{Q})$ is isomorphic to the dihedral group D_8. The group D_8 has a two-dimensional representation, which sends r to $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and s to $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. This can be visualized by thinking of D_8 as the group of symmetries of a square; cf. Figure 1. Thus we can construct a two-dimensional Artin representation

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

[2]For the purposes of this question it is reasonable to demand that ρ be ramified at only finitely many primes. See [Ram00] for a construction of an irreducible two-dimensional ρ which is ramified at infinitely many primes.

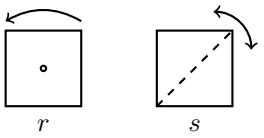


FIGURE 1. Generators for the dihedral group D_8, pictured as symmetries of the square.

which factors through $\mathrm{Gal}(\mathbf{Q}(i, \sqrt[4]{2})/\mathbf{Q})$. For an odd prime p, p is unramified in $\mathbf{Q}(i, \sqrt[4]{2})$, and we have the following explicit description of Frob_p:

$$\mathrm{Frob}_p \text{ is the conjugacy class of } \begin{cases} 1, & \text{if } p = a^2 + 64b^2, \\ r^2, & \text{if } p = a^2 + 16b^2 \text{ and } b \text{ is odd}, \\ rs, & \text{if } p \equiv 3 \bmod 8, \\ r, & \text{if } p \equiv 5 \bmod 8, \\ s, & \text{if } p \equiv 7 \bmod 8. \end{cases}$$

Thus in this situation we have a complete answer to Question C. Inasmuch as a representation of a group is determined by its *character*, the Galois representation ρ is determined by the function $\sigma \mapsto \mathrm{tr}\,\rho(\sigma)$. This function takes the following values on Frobenius elements:

$$\mathrm{tr}\,\rho(\mathrm{Frob}_p) = \begin{cases} 2, & \text{if } p = a^2 + 64b^2, \\ -2, & \text{if } p = a^2 + 16b^2,\ b \text{ odd}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4.1}$$

3.5. **p-adic Galois representations.** An Artin representation $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_n(\mathbf{C})$ always has a finite image. However, it is possible to construct Galois representations, even one-dimensional ones, having an infinite image. To do that, we change the target from a matrix group over $\mathbf{C}$ to one over the *p-adic numbers*. For a prime p, the *p-adic integers* $\mathbf{Z}_p$ are defined as the inverse limit

$$\mathbf{Z}_p = \varprojlim_m \mathbf{Z}/p^m\mathbf{Z}.$$

Alternatively, $\mathbf{Z}_p$ is the completion of $\mathbf{Z}$ with respect to the p-adic absolute value $n \mapsto |n|_p$, which for $n \neq 0$ is defined as the reciprocal of the largest power of p which divides n, and $|0|_p = 0$. A p-adic integer can be expressed uniquely as a series $a_0 + a_1 p + a_2 p^2 + \cdots$, with each "digit" $a_i \in \{0, 1, \ldots, p-1\}$. Let $\mathbf{Q}_p$ be the fraction field of $\mathbf{Z}_p$; this is the field of p-adic numbers.

A *p-adic Galois representation* has matrix entries in the field $\mathbf{Q}_p$ or a finite extension thereof. As a general rule, p-adic representations are far richer than Artin representations, because the topologies on $\mathrm{Gal}(\overline{K}/K)$ and $\mathrm{GL}_n(\mathbf{Q}_p)$ are more compatible. (The first group is profinite, and the second is locally profinite.)

At this point we caution the reader that when discussing p-adic Galois representations ρ, we will use the letter ℓ to denote a varying prime, so that for instance Question C will be about determining the conjugacy class of $\rho(\mathrm{Frob}_\ell)$ for unramified primes ℓ.

Example 3.5.1 (The p-adic cyclotomic character)**.** The p-adic cyclotomic character ρ_{cycl} is the one-dimensional representation

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}_p^\times \hookrightarrow \mathbf{Q}_p^\times = \mathrm{GL}_1(\mathbf{Q}_p)$$

defined as follows: if $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, then $\rho_{\mathrm{cycl}}(\sigma) = a_0 + a_1 p + a_2 p^2 + \cdots \in \mathbf{Z}_p^\times$ is characterized by the relation

$$\sigma(\zeta_{p^n}) = \zeta_{p^n}^{a_0 + a_1 p + \cdots + a_{n-1} p^{n-1}} \text{ for all } n,$$

where $\zeta_{p^n} = e^{2\pi i/p^n}$. The extensions $\mathbf{Q}(\zeta_{p^n})/\mathbf{Q}$ are ramified only at p [Was97, Proposition 2.3]. For a prime $\ell \neq p$, we have that $\mathrm{Frob}_\ell(\zeta_{p^n}) \equiv \zeta_{p^n}^\ell$ modulo a prime of $\mathbf{Q}(\zeta_{p^n})$ above ℓ, and this is enough to show that $\mathrm{Frob}_\ell(\zeta_{p^n}) = \zeta_{p^n}^\ell$ [Was97, Lemma 2.12]. It follows that $\rho_{\mathrm{cycl}}(\mathrm{Frob}_\ell) = \ell$, which answers Question C for ρ_{cycl}.

3.6. **Galois representations coming from geometry.** Algebraic varieties over number fields provide a rich source of Galois representations. As an example, let $f(x)$ be a monic polynomial of degree 3 with integer coefficients, and consider the plane curve $y^2 = f(x)$. Assume that $f(x)$ has no repeated roots, so that this curve is nonsingular. Let E be the completion of this curve, which is obtained by adding a point ∞. For any field K containing $\mathbf{Q}$, the set $E(K)$ of points of E with coordinates in K has the structure of an abelian group, with identity element ∞. In this group, three points sum to ∞ exactly when they are collinear.

The group of complex points $E(\mathbf{C})$ is isomorphic to a complex torus: $E(\mathbf{C}) \cong S^1 \times S^1$. For a positive integer n, the group $E[n]$ of n-torsion elements of $E(\mathbf{C})$ is therefore isomorphic to $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$. These "torsion points" all lie in $E(\overline{\mathbf{Q}})$. For instance, $E[2]$ consists of the origin ∞ together with the three points $(\alpha, 0)$, where α runs through the roots of $f(x)$. The group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts on $E(\overline{\mathbf{Q}})$ coordinate-wise and preserves $E[n]$. The action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $E[n]$ determines a Galois representation with values in $\mathrm{GL}_2(\mathbf{Z}/n\mathbf{Z})$.

Let p be a prime. Define the *p-adic Tate module*

$$T_p E = \varprojlim_m E[p^m] \cong \mathbf{Z}_p \times \mathbf{Z}_p = \mathbf{Z}_p^2.$$

The Tate module admits a continuous action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Viewing $\mathbf{Z}_p^2$ inside $\mathbf{Q}_p^2$, we obtain a Galois representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Q}_p)$. In contrast to the Artin representations, the image of ρ is infinite. It so happens that ρ is unramified at primes ℓ not dividing $p\Delta$, where Δ is the discriminant of $f(x)$. For such ℓ, the characteristic polynomial of the 2×2 matrix $\rho(\mathrm{Frob}_\ell)$ has the following shape:

$$\det(xI - \rho(\mathrm{Frob}_\ell)) = x^2 - (\ell + 1 - N_\ell)x + \ell, \tag{3.6.1}$$

where N_ℓ is the number of points of E with coordinates in the finite field $\mathbf{F}_\ell$.

More generally, let X be an algebraic variety defined over the rationals. For our purposes this is a system of polynomial equations with rational coefficients. For each $i \geq 0$ there is a finite-dimensional $\mathbf{Q}_p$-vector space $H^i(X_{\overline{\mathbf{Q}},\text{ét}}, \mathbf{Q}_p)$ (the ith *p-adic étale cohomology group*) admitting a continuous action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is unramified outside of a finite set of primes, namely p and those primes for which the reduction of X is singular.

Étale cohomology was introduced by Grothendieck in order to prove the Weil conjectures for a variety defined over a finite field. It is not quite the same as the usual (singular) cohomology of a topological space; the precise definition is

very technical.[3] Nonetheless, for a variety X, the étale cohomology groups are the same as the singular cohomology groups of the topological space $X(\mathbf{C})$. Thus when $X = E$ is our elliptic curve, so that $E(\mathbf{C}) \cong S^1 \times S^1$, we have $H^1(E_{\overline{\mathbf{Q}},\text{ét}}, \mathbf{Z}/n\mathbf{Z}) \cong H^1(S^1 \times S^1, \mathbf{Z}/n\mathbf{Z}) \cong \mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$. In fact $H^1(E_{\overline{\mathbf{Q}},\text{ét}}, \mathbf{Z}_p)$ is canonically isomorphic to the $\mathbf{Z}_p$-dual of the p-adic Tate module T_pE.

It is not much of an exaggeration to say that the only known constructions of Galois representations involve the étale cohomology of varieties in some way. (An Artin representation such as that appearing in Example 3.4.2 comes from the degree zero étale cohomology of a zero-dimensional variety!) There is a precise sense in which a Galois representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\mathbf{Q}_p)$ appears to come from geometry, or is "geometric" (we refer here to the property of being *potentially semi-stable* at p [Fon94]). The Fontaine–Mazur conjecture [FM95] asserts that an irreducible p-adic Galois representation which is geometric in this sense always appears as a subquotient of some $H^i(X_{\overline{\mathbf{Q}},\text{ét}}, \mathbf{Q}_p)$.

4. Elliptic modular forms

4.1. **Basic definition and examples.** The theory of modular forms developed in a context completely unrelated to the arithmetic questions posed in this article. They arose in relation to the elliptic functions investigated by Legendre, Abel, Jacobi, and others in the early 19th century, which in turn arose in association with finding the arc length of an ellipse. For an introduction to the elementary theory of modular forms, we recommend the book [Ser73].

A modular form is a certain kind of holomorphic function on the upper half-plane $\mathcal{H} = \{\tau | \mathrm{Im}\ \tau > 0\}$, which we view simultaneously as a complex manifold and as a Riemannian manifold equipped with the hyperbolic metric $y^{-2}(dx^2 + dy^2)$. The automorphism group of $\mathcal{H}$ is the group of *Möbius transformations* $\tau \mapsto \gamma \cdot \tau = (a\tau + b)/(c\tau + d)$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{R})$. In brief, a holomorphic function $f(\tau)$ on $\mathcal{H}$ is a modular form if it transforms in a certain way under a subgroup of $\mathrm{SL}_2(\mathbf{R})$.

Example 4.1.1. Before formally defining modular forms, we give an example. For $\tau \in \mathcal{H}$, the series

$$\theta(\tau) = \sum_{n \in \mathbf{Z}} e^{i\pi\tau n^2}$$

converges to a holomorphic function on $\mathcal{H}$. This is an example of a Jacobi theta function. Besides the evident relation $\theta(\tau+2) = \theta(\tau)$, it satisfies the transformation law

$$\theta(-1/\tau) = \sqrt{\tau/i}\theta(\tau), \tag{4.1.1}$$

[3] Here is a quick summary: If X is an algebraic variety, one has an *étale site* $X_{\text{ét}}$, which is something like a topological space without any points. The "open subsets" of $X_{\text{ét}}$ are not subsets at all but rather morphisms $U \to X$ that are *étale* (flat and unramified). Then one can define a sheaf $\mathcal{F}$ on the étale site: this is an assignment of a set $\mathcal{F}(U)$ to each $U \to X$, together with the appropriate restriction maps, which satisfies the sheaf axioms. From here it is more or less formal to define the cohomology $H^i(X_{\text{ét}}, \mathcal{F})$ whenever $\mathcal{F}$ is a sheaf of abelian groups. A special case is when $\mathcal{F} = \mathbf{Z}/n\mathbf{Z}$ is a constant sheaf. One defines $H^i(X_{\text{ét}}, \mathbf{Z}_p)$ as the inverse limit $\varprojlim H^i(X_{\text{ét}}, \mathbf{Z}/p^n\mathbf{Z})$, and $H^i(X_{\text{ét}}, \mathbf{Q}_p)$ as $H^i(X_{\text{ét}}, \mathbf{Z}_p) \otimes_{\mathbf{Z}_p} \mathbf{Q}_p$. The standard reference for étale cohomology is [Mil80]. See also the review [Blo81].

which can be proved using Poisson summation. The automorphisms $\tau \mapsto \tau + 2$ and $\tau \mapsto -1/\tau$ of $\mathcal{H}$ can be represented by the matrices $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, respectively. Thus θ admits a transformation law with respect to the subgroup of $\mathrm{SL}_2(\mathbf{Z})$ generated by those two matrices.

The modular form $\theta(\tau)$ is related to the Riemann zeta function $\zeta(s)$ via

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \int_0^\infty \frac{1}{2}(\theta(it) - 1)t^{\frac{s}{2}}\frac{dt}{t}. \tag{4.1.2}$$

Riemann's second proof of the analytic continuation and functional equation of $\zeta(s)$ relies on this relation. The main idea is to break up the integral in (4.1.2) into two pieces, one from 0 to 1 and the other from 1 to ∞. Then the transformation $t \mapsto 1/t$ is used on the former, using (4.1.1). The result is

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \int_1^\infty (t^{s/2} + t^{(1-s)/2})\frac{\theta(it) - 1}{2}\frac{dt}{t} - \frac{1}{s} - \frac{1}{1-s}.$$

Since $\theta(it) - 1 = O(e^{-\pi t})$ as $t \to \infty$, $\pi^{-s/2}\Gamma(s/2)\zeta(s)$ extends to a meromorphic function on $\mathbf{C}$ invariant under $s \mapsto 1 - s$.

We will only define modular forms associated to subgroups of $\mathrm{SL}_2(\mathbf{R})$ of a very particular shape. For a nonzero integer N, let

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) \;\middle|\; a, d \equiv 1 \bmod N,\ c \equiv 0 \bmod N \right\}. \tag{4.1.3}$$

This is a subgroup of $\mathrm{SL}_2(\mathbf{Z})$ of finite index.

Definition 4.1.2. Let $N, k \geq 1$ be integers. A *modular form of weight k and level N* is a holomorphic function g on $\mathcal{H}$ which satisfies

$$g\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k g(\tau)$$

for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(N)$, and which is "holomorphic at the cusps".

Since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_1(N)$, a modular form $g(\tau)$ satisfies $g(\tau + 1) = g(\tau)$, so that g is a function of the parameter $q = e^{2\pi i\tau}$. "Holomorphic at the cusps" means that the Fourier expansion of $g(\tau)$, a priori a series of the form $\sum_{n \in \mathbf{Z}} a_n(g)q^n$, has $a_n(g) = 0$ for $n < 0$; a similar condition is imposed for $g \circ \gamma$ for all $\gamma \in \mathrm{SL}_2(\mathbf{Z})$. We say g is a *cusp form* (or that g is *cuspidal*) if it is zero at the cusps, meaning that $a_0(g \circ \gamma) = 0$ as well.

If g is a modular form of weight k, we set $L(g, s) = \sum_{n \geq 1} a_n(g)n^{-s}$, which turns out to be convergent for $\mathrm{Re}\ s > k$. A manipulation along the lines of Example 4.1.1 shows that $L(g, s)$ admits an analytic continuation to all of $\mathbf{C}$ (entire if g is a cusp form) and satisfies a functional equation relating $L(g, k - s)$ to $L(\widetilde{g}, s)$, where $\widetilde{g}(\tau) = g(-1/(N\tau))$.

Example 4.1.3 (Eisenstein series)**.** Let $k \geq 4$ be even. For $\tau \in \mathcal{H}$, the series

$$G_k(\tau) = \sum_{(m,n) \in \mathbf{Z}^2 \setminus \{(0,0)\}} \frac{1}{(m\tau + n)^k}$$

defines a holomorphic function of τ, which turns out to be a modular form of weight k and level 1. That is, it satisfies

$$G_k\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k G_k(\tau)$$

for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(1) = \mathrm{SL}_2(\mathbf{Z})$. Its Fourier expansion is

$$G_k(\tau) = 2\zeta(k)\left(1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n\right),$$

where $q = e^{2\pi i\tau}$, $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$ and $B_k \in \mathbf{Q}$ is the kth Bernoulli number. If we normalize the Eisenstein series by setting $E_k(\tau) = G_k(\tau)/(2\zeta(k))$, then the Fourier expansion of $E_k(\tau)$ has rational coefficients and constant term 1. Then E_k is a modular form of weight k and level 1, and $L(E_k, s) = \zeta(s)\zeta(s-k+1)$.

Example 4.1.4 (The cusp form Δ)**.** The Eisenstein series E_k are not cusp forms, since their Fourier expansions have nonzero constant term. Define $\Delta(\tau) = (E_4(\tau)^3 - E_6(\tau)^2)/1728$; this is a cusp form of weight 12 and level 1. It has integral Fourier coefficients:

$$\Delta(\tau) = q\prod_{n\geq 1}(1-q^n)^{24} = \sum_{n\geq 1}\tau(n)q^n,$$

where the sequence[4] $\tau(n)$ for $n \geq 1$ starts out as $1, -24, 252, -1472, \ldots$. It was observed by Ramanujan and proved by Mordell that $\tau(mn) = \tau(m)\tau(n)$ when m and n are relatively prime [Ser73, Cor. to Prop. 14].

4.2. **Some modular forms of weight 1.** If θ is the Jacobi theta function of Example 4.1.1, then

$$\theta(2\tau)^2 = \sum_{a,b\in\mathbf{Z}} q^{a^2+b^2}$$

is a modular form of weight 1 and level 4.

This is an instance of a very general construction involving rings of integers in quadratic fields, such as $\mathbf{Z}[i]$. Suppose $\alpha \subset \mathbf{Z}[i]$ is nonzero and $\chi\colon (\mathbf{Z}[i]/(\alpha))^\times \to \mathbf{C}^\times$ is a homomorphism. Assume that $\chi(i) = 1$. Extend χ to a multiplicative function on $\mathbf{Z}[i]$ by declaring it to be 0 on elements which are not prime to α. It is a result of Hecke [Hec27] that the series

$$\theta_\chi(\tau) = \frac{1}{4}\sum_{a,b\in\mathbf{Z}}\chi(a+bi)q^{a^2+b^2}$$

is a modular form of weight 1 and level $4\left|\alpha\right|^2$, and if χ is nontrivial, then θ_χ is a cusp form. (Note that if $\chi(i) \neq 1$, this series is 0.)

Example 4.2.1 (A cuspidal theta function)**.** The abelian group $(\mathbf{Z}[i]/8\mathbf{Z}[i])^\times$ has generators 3, 5, i, and $1+2i$, of orders 2, 2, 4, and 4, respectively. Let $\chi\colon (\mathbf{Z}[i]/8\mathbf{Z}[i])^\times \to \mathbf{C}^\times$ be the unique homomorphism which is trivial on the first three generators and which sends $1+2i$ to i. Then θ_χ is a modular form of weight 1 and level 256. For a prime p, the pth coefficient in the Fourier expansion of θ_χ is

$$a_p(\theta_\chi) = \begin{cases} \chi(a+bi) + \chi(a-bi), & p \equiv 1 \bmod 4,\ p = a^2+b^2, \\ 0, & p \equiv 3 \bmod 4 \text{ or } p = 2. \end{cases}$$

[4]The τ in $\tau(n)$ has nothing to do with the complex variable we write as τ. Both notations are traditional.

Now if $p \equiv 1 \pmod 4$, we can write $p = a^2 + b^2$ with a odd and b even. A short calculation shows that

$$a_p(\theta_\chi) = \begin{cases} 2, & 8|b, \\ -2, & 4|b \text{ but } 8 \nmid b, \\ 0, & 4 \nmid b. \end{cases}$$

Referring back to (3.4.1), we find that for all odd primes p,

$$a_p(\theta_\chi) = \operatorname{tr} \rho(\mathrm{Frob}_p)$$

for the Galois representation ρ constructed in Example 3.4.2. This equation hints at an extraordinary relationship between modular forms and representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

4.3. **Hecke operators.** Let $M_k(N)$ (resp., $S_k(N)$) denote the space of modular forms (resp., cusp forms) of weight k and level N. Then $M_k(N)$ and $S_k(N)$ are finite-dimensional vector spaces over $\mathbf{C}$. For each prime p not dividing N, the *Hecke operators* T_p and $\langle p \rangle$ are endomorphisms of $M_k(N)$ defined by

$$\begin{aligned} (T_p g)(\tau) &= p^{k-1} g(p\tau) + \frac{1}{p} \sum_{j=0}^{p-1} g\left(\frac{\tau + j}{p} \right), \\ (\langle p \rangle g)(\tau) &= g(\gamma\tau)(c\tau + d)^{-k}. \end{aligned}$$

Here $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is any element of $\mathrm{SL}_2(\mathbf{Z})$ with $c \equiv 0 \pmod N$ and $d \equiv p \pmod N$. There are similar operators T_n and $\langle n \rangle$ for all integers n which are relatively prime to N. The Hecke operators preserve $M_k(N)$ and $S_k(N)$.

These operators commute with one another. Furthermore there is a Hermitian inner product on $S_k(N)$ relative to which these operators are normal. Therefore the T_n and $\langle n \rangle$ are simultaneously diagonalizable on $S_k(N)$. A modular form is an *eigenform* if it is an eigenvector for all the Hecke operators.[5] Suppose $g = \sum_{n \geq 1} a_n(g) q^n$ is a cuspidal eigenform which is *normalized*, meaning that $a_1(g) = 1$. Then for all n relatively prime to N, the eigenvalue of T_n on g is just $a_n(g)$.

If $g(\tau) \in M_k(N)$ is an eigenform, then there exists a homomorphism $\chi : (\mathbf{Z}/N\mathbf{Z})^\times \to \mathbf{C}^\times$ (a *Dirichlet character*) such that $\langle n \rangle g(\tau) = \chi(n) g(\tau)$ for all n relatively prime to N. Then g satisfies

$$g\left(\frac{a\tau + b}{c\tau + d} \right) = \chi(d)(c\tau + d)^k g(\tau)$$

for all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ lying in the group

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) \;\middle|\; c \equiv 0 \bmod N \right\}.$$

We say that χ is the character of $g(\tau)$.

It so happens that the space $S_{12}(1)$ is one dimensional, spanned by Δ. Thus it must necessarily be an eigenform for all the Hecke operators T_n, with eigenvalue $\tau(n)$. This explains why $\tau(mn) = \tau(m)\tau(n)$ when m and n are relatively prime.

[5]There are also Hecke operators indexed by primes which do divide N; an eigenform should be an eigenvector for these also.

Let $g(\tau) \in S_k(N)$ have Fourier expansion $\sum_{n\geq 1} a_n q^n$. Assume that $a_1 = 1$. It is formal to show that if $g(\tau)$ is an eigenform with character χ, then $L(g,s)$ can be written as an Euler product

$$L(g,s) = \prod_p L_p(g,s),$$

where $L_p(g,s)$ is the reciprocal of a polynomial in p^{-s} of degree at most 2. For every p not dividing N,

$$L_p(g,s) = \frac{1}{1 - a_p p^{-s} + \chi(p)p^{k-2s}}.$$

See for instance [Kob84, Prop. 36].

Remarkably, if $g(\tau)$ is an eigenform, then there exists a number field F for which the Hecke eigenvalues of $g(\tau)$ all lie in $\mathcal{O}_F$.

4.4. **Galois representations associated with modular forms.** The following theorem of Deligne and Serre generalizes the phenomenon we observed in Example 4.2.1. For a number field F and a prime $\mathfrak{p}$ of F, we let $F_\mathfrak{p}$ denote the completion of F with respect to the $\mathfrak{p}$-adic absolute value. It is a finite extension of a p-adic field $\mathbf{Q}_p$. We remark here that a two-dimensional Galois representation ρ is *even* or *odd* as $\det \rho(c)$ is 1 or -1, respectively, where c is complex conjugation.

Theorem 4.4.1. *Let $g(\tau) = \sum_{n\geq 1} a_n(g) q^n$ be a cuspidal eigenform of weight k, level N, and character $\chi\colon (\mathbf{Z}/N\mathbf{Z})^\times \to \mathbf{C}^\times$, normalized so that $a_1 = 1$. Let F be a number field containing the $a_n(g)$ and the values of χ.*

1. *Suppose $k \geq 2$. Then for all primes $\mathfrak{p}$ of F there exists an odd irreducible Galois representation*

$$\rho_{g,\mathfrak{p}}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(F_\mathfrak{p})$$

   such that for all ℓ prime to N and to $\mathfrak{p}$, $\rho_{g,\mathfrak{p}}$ is unramified at ℓ, and the characteristic polynomial of $\rho_{g,\mathfrak{p}}(\mathrm{Frob}_\ell)$ is $x^2 - a_\ell(g)x + \chi(\ell)\ell^{k-1}$.
2. *Suppose $k = 1$. Then there exists an odd irreducible Galois representation*

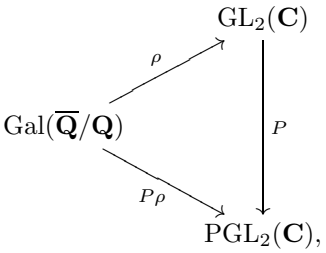
$$\rho_g\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$$

   such that for all ℓ prime to N, ρ_g is unramified at ℓ, and the characteristic polynomial of $\rho_g(\mathrm{Frob}_\ell)$ is $x^2 - a_\ell(g)x + \chi(\ell)$.

These two statements are proved in [Del71] and [DS74], respectively. In the first statement, the image of $\rho_{g,\mathfrak{p}}$ is infinite. In the second statement, ρ_g is an Artin representation whose image is finite.

Example 4.4.2 (An icosahedral cusp form)**.** Let $f(x)$ be a polynomial of degree 5 with rational coefficients, and let K be its splitting field. Then $\mathrm{Gal}(K/\mathbf{Q})$ is a subgroup of S_5, the group of permutations of the five roots of $f(x)$. Let us assume that $\mathrm{Gal}(K/\mathbf{Q})$ is isomorphic to A_5, the group of even permutations. The group A_5 does not have any irreducible two-dimensional representations, but there exists an extension $\widetilde{A}_5$ of A_5 by the cyclic group of order 4 which does. It can be shown that there is an irreducible Artin representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ whose

image is isomorphic to $\widetilde{A}_5$, such that in the diagram



where P is the projection homomorphism, the fixed field of the kernel of $P\rho$ is K. The Artin representation ρ is odd when $P\rho(c) \neq 1$ (c being complex conjugation), which is equivalent to the condition that c act nontrivially on the roots of $f(x)$. Assume this is the case. Then since the action of c on the roots of $f(x)$ is an even permutation of order 2, it must be the product of two transpositions. We deduce that $f(x)$ has exactly one real root.

Recall that Theorem 4.4.1 associates an odd irreducible Artin representation ρ_g to a cuspidal eigenform g of weight 1. Does $\rho = \rho_g$ for such an eigenform g? This question was answered affirmatively by Buhler [Buh78] for the polynomial

$$f(x) = x^5 + 10x^3 - 10x^2 + 35x - 18.$$

In this case $\rho = \rho_g$, where g is a cuspidal eigenform

$$g(\tau) = q - iq^3 - ijq^7 - q^9 + jq^{13} + i(1-j)q^{19} - jq^{21} + \cdots,$$

where $i = \sqrt{-1}$ and $j = (1+\sqrt{5})/2$. This g belongs to $S_1(800)$ and its character χ factors through a homomorphism $(\mathbf{Z}/100\mathbf{Z})^\times \to \mathbf{C}^\times$. A prime $\ell \neq 2, 5$ splits in K if and only if $\rho(\mathrm{Frob}_\ell)$ is a scalar matrix. Since $\rho(\mathrm{Frob}_\ell)$ has finite order, it is semisimple, and therefore it is scalar if and only if its characteristic polynomial has zero discriminant. The characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is $x^2 - a_\ell(g)x + \chi(\ell)$, with discriminant $a_\ell(g)^2 - 4\chi(\ell)$. Therefore we have the following answer to Question B: ℓ splits in K if and only if $a_\ell(g)^2 = 4\chi(\ell)$.

Example 4.4.3 (The Galois representation associated to Δ)**.** Recall from Example 4.1.4 the cuspidal eigenform $\Delta(\tau) = \sum_{n\geq 1} \tau(n)q^n$ of weight 12 and level 1 (with trivial character). Theorem 4.4.1 associates to Δ a p-adic representation $\rho_{\Delta,p}$ for all primes p. This can be reduced modulo p to obtain a mod p Galois representation $\overline{\rho}_{\Delta,p}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Z}/p\mathbf{Z})$, whose kernel cuts out a number field which is ramified only at p. It is a difficult computational problem to compute this number field. For some small primes p this has been carried out in [Bos11], at least for the associated projective representation $P\overline{\rho}_{\Delta,p}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{PGL}_2(\mathbf{Z}/p\mathbf{Z})$. For instance if $p = 11$, the fixed field of the kernel of $P\overline{\rho}_{\Delta,p}$ is the splitting field of

$$\begin{aligned} f(x) \;=\; & x^{12} - 4x^{11} + 55x^9 - 165x^8 + 264x^7 - 341x^6 \\ & + 330x^5 - 165x^4 - 55x^3 + 99x^2 - 41x - 111. \end{aligned}$$

From this we can derive a partial answer to Question A for this $f(x)$, valid for almost all primes[6] ℓ: if $f(x)$ splits modulo ℓ, then $P\overline{\rho}_{\Delta,11}(\mathrm{Frob}_\ell)$ is the identity, so $\overline{\rho}_{\Delta,11}$ is a scalar aI_2, where $a \in \mathbf{Z}/11\mathbf{Z}$. The characteristic polynomial of $\rho_\Delta(\mathrm{Frob}_\ell)$ is $x^2 - \tau(\ell)x + \ell^{11}$. Therefore $2a \equiv \tau(\ell) \pmod{11}$ and $a^2 \equiv \ell^{11} \equiv \ell \pmod{11}$, and

[6] "Almost all primes" means "all but finitely many primes".

so we have $\tau(\ell)^2 \equiv 4\ell \pmod{11}$. This is an interesting necessary condition for $f(x)$ to split mod ℓ. Unfortunately the converse is not true: if $\tau(\ell)^2 \equiv 4\ell \pmod{11}$, the most we can say about $\rho_{\Delta,11}(\mathrm{Frob}_\ell)$ is that it is conjugate to a matrix of the form $\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right) \pmod{11}$.

4.5. **Modular Galois representations.** Theorem 4.4.1 says that to every cuspidal eigenform g and prime number p there exists an associated two-dimensional odd irreducible p-adic Galois representation ρ, such that ρ is unramified at ℓ for almost all primes ℓ, and the conjugacy class of $\rho(\mathrm{Frob}_\ell)$ can be read off from the eigenvalues of T_ℓ and $\langle \ell \rangle$ on g. Let us call an odd irreducible Galois representation *modular* if it arises this way. Thus we have an affirmative answer to Question C for all modular Galois representations.

The question of which Galois representations are modular took on a special urgency in the 1980s, because of the remarkable link to Fermat's Last Theorem. Recall from §3.6 that if E is an elliptic curve over the rational numbers and p is a prime, then the Tate module T_pE is a two-dimensional p-adic Galois representation. We say that E is *modular* if T_pE is. Thus E is modular if and only if there exists a cuspidal eigenform g of weight k and character χ such that for almost all primes ℓ, the characteristic polynomial of Frob_ℓ on T_pE is $x^2 - a_\ell(g)x + \chi(\ell)\ell^{k-1}$. Comparing this with (3.6.1), we find that E is modular if and only if there exists a cuspidal eigenform g of weight 2 and trivial character, such that for almost all primes ℓ, the number of points on E with coordinates in $\mathbf{F}_\ell$ is $\ell + 1 - a_\ell(g)$. Note that this statement is independent of the prime p!

Conjecture 4.5.1 (The Shimura–Taniyama–Weil conjecture)**.** *Every elliptic curve defined over the rational numbers is modular.*

Suppose $p \geq 3$ is prime and a, b, c are nonzero integers with $a^p + b^p = c^p$. It was pointed out by Frey [Fre86] that (possibly after normalizing the triple (a, b, c) a little bit) the elliptic curve E with equation $y^2 = x(x - a^p)(x + b^p)$ has unusual properties suggesting that it could not be modular. These properties are related to the fact that the discriminant of the cubic $x(x - a^p)(x + b^p)$ is $(abc)^{2p}$, a perfect pth power. Ribet [Rib90] showed that E is not modular, using subtle properties of modular forms. The next breakthrough came with Wiles [Wil95] and Taylor and Wiles [TW95], who showed that E is modular after all! With this contradiction fell Fermat's Last Theorem. In fact [TW95] proved the modularity of any *semistable* elliptic curve, a class of elliptic curves containing Frey's curve. The full Shimura–Taniyama–Weil conjecture (applying to all elliptic curves over $\mathbf{Q}$) was proved in [BCDT01].

The Tate module of an elliptic curve gives an odd irreducible p-adic representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Q}_p)$. Could it be the case that every odd irreducible p-adic representation ρ is modular? Not quite, because there are a few necessary conditions. One is that ρ be unramified at all but finitely many primes. There is also the "geometric" condition we mentioned at the end of §3.6. A modular representation satisfies both conditions.

Conjecture 4.5.2 (Fontaine and Mazur [FM95])**.** *Let $F/\mathbf{Q}_p$ be a finite extension, and let $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(F)$ be an odd irreducible geometric Galois representation which is unramified at all but finitely many primes. Then ρ is modular.*

Many cases of this conjecture were proved independently by Emerton [Eme06] and Kisin [Kis09].

If $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ is an irreducible odd *Artin* representation, then ρ is modular if and only if there exists a cuspidal eigenform g of weight 1 such that $\operatorname{tr}\rho(\mathrm{Frob}_p) = a_p(g)$ for almost all primes p. This is the two-dimensional case of *Artin's conjecture*, which can be stated for all dimensions in terms of the analytic continuation of an L-function attached to ρ. (The one-dimensional case of Artin's conjecture is settled by class field theory.) Two-dimensional Artin representations can be classified by the projective image of ρ, which is a finite subgroup in $\mathrm{PGL}_2(\mathbf{C})$; this can be dihedral ($D_n$), tetrahedral ($A_4$), octahedral ($S_4$), or icosahedral ($A_5$). In the dihedral case, the required eigenform g is a theta function, similar to the one appearing in Example 4.2.1. The tetrahedral and octahedral cases were treated by Langlands [Lan80] and Tunnell [Tun81], respectively, using an analytic technique known as *base change*; these results were used in Wiles's attack on Fermat's Last Theorem. Much more difficult is the icosahedral case, essentially because A_5 is nonsolvable. All cases of Artin's conjecture for two-dimensional odd Artin representations were finally settled in [KW09b]. Thus Question A is settled for polynomials $f(x)$ of degree ≤ 5 with Galois group A_5 having exactly one real root, as in Example 4.4.2.

4.6. **Modular curves.** Here we sketch out some of the ideas behind the proof of Theorem 4.4.1, concerning the existence of Galois representations attached to cuspidal eigenforms.

Modular forms are holomorphic functions on $\mathcal{H}$ which admit symmetries with respect to a finite-index subgroup $\Gamma \subset \mathrm{SL}_2(\mathbf{Z})$. It stands to reason that they correspond to objects defined on the quotient $\Gamma\backslash\mathcal{H}$, a (noncompact) Riemann surface. To illustrate this, suppose that $g(\tau)$ is a modular form of weight 2 and level N. Then the differential form $g(\tau)d\tau$ on $\mathcal{H}$ is invariant under $\Gamma_1(N)$ and so descends to a differential form on $\Gamma_1(N)\backslash\mathcal{H}$.

The curve $\Gamma_1(N)\backslash\mathcal{H}$ can be compactified by adding a finite set of points called *cusps*, one for each element of $\Gamma_1(N)\backslash\mathbf{P}^1(\mathbf{Q})$. The result is a compact Riemann surface called $X_1(N)$, whose underlying set of points is $\Gamma_1(N)\backslash(\mathcal{H}\cup\mathbf{P}^1(\mathbf{Q}))$. If $g(\tau)$ happens to be a cusp form, then $g(\tau)d\tau$ extends to a differential form on $X_1(N)$. In fact the space $S_2(N)$ of cusp forms of weight 2 and level N is isomorphic to the space $H^0(X_1(N), \Omega^1_{X_1(N)/\mathbf{C}})$ of (holomorphic) differential forms on $X_1(N)$.

The Riemann surfaces $X_1(N)$ are examples of *modular curves*. Modular curves are defined in general as those Riemann surfaces arising as the quotient of $\mathcal{H}$ by a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$ (that is, a subgroup defined by congruence conditions modulo N) along with their compactifications. Any hyperbolic Riemann surface is a quotient of $\mathcal{H}$ by some discrete subgroup $\Gamma \subset \mathrm{SL}_2(\mathbf{R})$, but modular curves are distinguished by the condition that Γ be conjugate to a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$.

Modular curves come equipped with a family of multivalued functions known as *Hecke correspondences*. For an element $\tau \in \mathcal{H}\cup\mathbf{P}^1(\mathbf{Q})$, let $[\tau]$ denote its image in $X_1(N)$. If $\gamma \in \mathrm{GL}_2(\mathbf{Q})$ has positive determinant, then $[\tau] \mapsto [\gamma(\tau)]$ is not generally a well-defined function $X_1(N) \to X_1(N)$, but it only takes finitely many values. This is the Hecke correspondence associated to γ; it only depends on the double coset $\Gamma_1(N)\gamma\Gamma_1(N)$. For each prime $p \nmid N$, the Hecke correspondence associated to the matrix $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is denoted T_p, and the Hecke correspondence associated to $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$ ($c \equiv 0 \pmod N$, $d \equiv p \pmod N$) is denoted $\langle p\rangle$. Each Hecke correspondence determines an endomorphism of the cohomology group $H^1(X_1(N), \mathbf{Q})$.

We can now sketch a proof of Theorem 4.4.1 in the case of weight 2, which synthesizes a study of modular curves from the analytic and algebraic points of view. On the analytic side we have the Hodge decomposition for the compact Riemann surface $X_1(N)$:

$$H^1(X_1(N), \mathbf{Q}) \otimes_{\mathbf{Q}} \mathbf{C} \cong H^{1,0} \oplus H^{0,1},$$

where $H^{1,0} = H^0(X_1(N), \Omega^1_{X_1(N)/\mathbf{C}}) \cong S_2(N)$ and $H^{0,1} = \overline{H}^{1,0}$. The spaces $H^1(X_1(N), \mathbf{Q})$ and $S_2(N)$ come equipped with actions by the Hecke operators T_p and $\langle p \rangle$ for $p \nmid N$, and the isomorphism above is compatible with the action of these operators. The conclusion is that systems of eigenvalues coming from cuspidal eigenforms of weight 2 and level N appear in $H^1(X_1(N), \mathbf{C})$ with multiplicity 2. (There is a similar statement for forms of higher weight; one replaces the $\mathbf{C}$ in $H^1(X_1(N), \mathbf{C})$ with a nonconstant coefficient system.) Since $H^1(X_1(N), \mathbf{Q})$ is a finite-dimensional vector space over $\mathbf{Q}$, the eigenvalues of the Hecke operators T_p and $\langle p \rangle$ all belong to a single number field, and therefore the same is true for Hecke eigenvalues of cusp forms.

On the algebraic side, we have the remarkable fact that $X_1(N)$ is a projective algebraic curve *which can be defined over the rational numbers*. All compact Riemann surfaces are projective algebraic curves—this is Riemann's existence theorem. Thus $X_1(N)$ can be realized as the set of solutions to a system of polynomial equations. Much harder is the statement that the coefficients of these polynomials can be taken to be rational numbers. This is related to the interpretation of modular curves as *moduli spaces for elliptic curves with level structure*. In this interpretation, the point $[\tau]$ of $Y_1(N)$ corresponds to the elliptic curve $\mathbf{C}/(\mathbf{Z} \oplus \mathbf{Z}\tau)$ together with the N-torsion point $1/N$. The problem of classifying elliptic curves E together with a point of order N can be posed over the rational numbers, and its solution is a rational model for $Y_1(N)$. With some care $Y_1(N)$ can be defined as the solution to a moduli problem over the integers, which is representable by a scheme over $\operatorname{Spec} \mathbf{Z}$ which is smooth modulo p for $p \nmid N$. The definitive reference for this topic is [KM85].

As a result, we can form the étale cohomology $H^1(X_1(N)_{\text{ét}}, \mathbf{Q}_p)$ as in §3.6; this is a representation of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is unramified outside of pN. Furthermore, the Hecke correspondences are also defined over the rational numbers, which means that their action on $H^1(X_1(N)_{\text{ét}}, \mathbf{Q}_p)$ commutes with the action of Galois.

The analytic and algebraic stories are linked together by means of the following comparison isomorphism:

$$H^1(X_1(N), \mathbf{Q}) \otimes \mathbf{Q}_p \cong H^1(X_1(N)_{\text{ét}}, \mathbf{Q}_p). \tag{4.6.1}$$

Suppose now that g is a cuspidal eigenform of weight 2 and level N. For simplicity assume that the Hecke eigenvalues of g lie in $\mathbf{Q}$. This means that there exists a two-dimensional Hecke eigenspace $V \subset H^1(X_1(N), \mathbf{Q})$ with the same Hecke eigenvalues as g. Let V_p be the image of $V \otimes_{\mathbf{Q}} \mathbf{Q}_p$ under (4.6.1). Since the actions of the Hecke operators and $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ commute, the latter preserves V_p. The action of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the dual of V_p is the Galois representation $\rho_{g,p}$ required by Theorem 4.4.1. One still needs to verify that for all primes ℓ not dividing pN, the characteristic polynomial of $\rho_{g,p}(\operatorname{Frob}_\ell)$ is as claimed. This is a consequence of the *Eichler–Shimura relation*, for which we refer the reader to [PS73].

4.7. **Interlude: Automorphic representations.** This article began with a discussion of the problem of determining the splitting behavior of a polynomial modulo a prime (Question A). This was refined into a question about the splitting behavior of primes in extensions of number fields (Question B). A further refinement posed the same problem in terms of Galois representations (Question C): given a Galois representation $\rho\colon \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_n(F)$, is there a rule for determining the conjugacy class of $\rho(\mathrm{Frob}_{\mathfrak{p}})$ for the unramified primes $\mathfrak{p}$?

Here is a summary of what we have discussed so far concerning Question C:

- The case $n = 1$ finds a satisfactory solution in class field theory. For instance if $K = \mathbf{Q}$ and $F = \mathbf{C}$, characters $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{C}^\times$ are essentially in correspondence with Dirichlet characters χ; the correspondence is characterized by $\rho(\mathrm{Frob}_\ell) = \chi(\ell)$ for almost every ℓ, as in Example 3.4.1.
- In the case $n = 2$ and $K = \mathbf{Q}$, a large class of p-adic Galois representations are *modular*, which means there exists a corresponding cuspidal eigenform g of some weight k, level N, and character χ on $(\mathbf{Z}/N\mathbf{Z})^\times$. Then for every prime $\ell \nmid Np$, the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is $x^2 - a_\ell(g)x + \chi(\ell)\ell^{k-1}$.

Dirichlet characters and eigenforms, though they seem like completely different entities, are but two instances of a class of objects called *automorphic representations*, which belong to the realm of harmonic analysis. The study of automorphic representations and their relation to Galois groups is known as the *Langlands program*, after Robert Langlands, who laid out a series of sweeping conjectures unifying the phenomena described above. See [Gel84] for a detailed introduction. Before saying more, let us highlight some of the properties shared by Dirichlet characters and eigenforms:

- **Local data.** A Dirichlet character χ can be described in terms of local data $\chi(p)$ for every prime p. Similarly, an eigenform g can be described in terms of its eigenvalues $a_p(g)$ and $\chi(p)$ for the Hecke operators T_p and $\langle p \rangle$, respectively.
- **L-functions.** A Dirichlet character χ has an L-function $L(\chi, s)$, which factors as an Euler product $\prod_p (1 - \chi(p)p^{-s})^{-1}$ for $\mathrm{Re}\ s > 1$, where the pth Euler factor depends on the local data of χ at p. Similarly, an eigenform g has an L-function $L(g, s)$, which factors as an Euler product $\prod_p (1 - a_p(g)p^{-s} + \chi(p)p^{k-2s})^{-1}$ for $\mathrm{Re\ s} > k$. Both $L(\chi, s)$ and $L(g, s)$ admit an analytic continuation and functional equation. In fact $L(\chi, s)$ (resp., $L(g, s)$) is entire if χ is nontrivial (resp., if g is cuspidal).
- **Galois representations.** Both Dirichlet characters and eigenforms have associated Galois representations ρ, of dimensions 1 and 2 respectively, where for almost all primes p, the conjugacy class of $\rho(\mathrm{Frob}_p)$ is determined by the local data at p of the character or eigenform.

An automorphic representation, then, must be some sort of entity π which has local components π_p for all primes p, as well as an L-function $L(\pi, s) = \prod_p L(\pi_p, s)$, initially convergent for $\mathrm{Re}\ s \gg 0$, but which has an analytic continuation and functional equation. For some automorphic representations (the *algebraic* ones), we expect there exists a corresponding Galois representation ρ, for which the conjugacy class $\rho(\mathrm{Frob}_p)$ is determined by π_p.

We had mentioned that automorphic representations belong to the realm of harmonic analysis. In harmonic analysis one starts with a measure space X and considers the Hilbert space $L^2(X)$ of square-integrable[7] functions on X. If X has a measure-preserving right action by a group G, then $L^2(X)$ becomes a unitary representation of G, via $(g \cdot f)(x) = f(xg)$, for $f \in L^2(X)$, $x \in X$, $g \in G$. We are interested in the decomposition of $L^2(X)$ into irreducible unitary G-modules. For instance if $G = X = \mathbf{R}/\mathbf{Z}$ (acting on itself by addition), then the irreducible unitary $\mathbf{R}/\mathbf{Z}$-modules appearing in $L^2(\mathbf{R}/\mathbf{Z})$ are one-dimensional and are spanned by the functions $e_n(z) = e^{2\pi inz}$, $n \in \mathbf{Z}$. Any function f in $L^2(\mathbf{R}/\mathbf{Z})$ has a Fourier expansion

$$f(z) = \sum_{n \in \mathbf{Z}} c_n(f) e_n(z),$$

and so $L^2(\mathbf{R}/\mathbf{Z})$ decomposes as a Hilbert direct sum of the one-dimensional irreducible representations of $\mathbf{R}/\mathbf{Z}$. If $X = G = \mathbf{R}$, the situation is more subtle: the unitary irreducible representations of $\mathbf{R}$ are the characters $e_t(z) = e^{2\pi itz}$, $t \in \mathbf{R}$, but these do not belong to $L^2(\mathbf{R})$. Nevertheless, every $f \in L^2(\mathbf{R})$ can be expressed as a Fourier transform:

$$f(z) = \int_{t \in \mathbf{R}} \widehat{f}(t) e_t(z) dt.$$

One says that the representations e_t appear continuously in $L^2(\mathbf{R})$.

Now let $G = \mathrm{SL}_2(\mathbf{R})$, and let $\Gamma \subset G$ be a discrete subgroup of finite covolume. (For instance, Γ could be a subgroup of $\mathrm{SL}_2(\mathbf{Z})$ of finite index.) The decomposition of $L^2(\Gamma\backslash G)$ into unitary representations of G leads naturally to modular forms on Γ. For the full story, see [Gel75, §2]. The basic idea is that $L^2(\Gamma\backslash G)$ has a discrete series part coming from modular forms and (if $\Gamma\backslash G$ is not compact) a continuous series part coming from Eisenstein series. If g is a modular form of weight k for Γ, we can define a function $\phi_g \in L^2(\Gamma \backslash \mathrm{SL}_2(\mathbf{R}))$, via

$$\phi_g(\gamma) = (ci + d)^{-k} g\left(\frac{ai + b}{ci + d}\right), \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{4.7.1}$$

and then the G-module generated by ϕ_g is a so-called discrete series representation of weight k. In fact there is a lovely decomposition of the discrete series part of $L^2(\Gamma\backslash G)$ into irreducible subspaces corresponding to modular forms (along with their nonholomorphic cousins, the Maass forms).

Thus it is tempting to define a general automorphic representation as a unitary representation of a Lie group G appearing in $L^2(\Gamma\backslash G)$, where $\Gamma \subset G$ is a discrete subgroup. But this won't quite suffice, because it is not clear how the required local data at every prime p are going to appear.

The correct formalism is going to put all the prime numbers on an equal footing with the infinite place of $\mathbf{Q}$, and for this we need the *adele ring* $\mathbf{A}$. Recall that for each prime p, we have the field of p-adic numbers $\mathbf{Q}_p$. Define also $\mathbf{Q}_\infty = \mathbf{R}$, the completion of $\mathbf{Q}$ with respect to the usual metric. We define $\mathbf{A}$ as the subring of $\prod_{p \leq \infty} \mathbf{Q}_p$ consisting of those $(a_p)_{p \leq \infty}$ for which $a_p \in \mathbf{Z}_p$ for almost all p. Note that $\mathbf{A}$ contains $\mathbf{Q}$ as a subring, embedded diagonally.

Let $n \geq 1$. The focus now shifts to the group $\mathrm{GL}_n(\mathbf{A})$. An appropriate choice of topology on $\mathrm{GL}_n(\mathbf{A})$ gives it the structure of a nondiscrete locally compact

[7]Pedantic note: In practice one usually modifies this to "square-integrable modulo center" after selecting a central character, in order for the theory to work properly.

topological group. (This is not the topology induced by viewing $\mathrm{GL}_n(\mathbf{A})$ as a subgroup of $\prod_{p\leq\infty}\mathrm{GL}_n(\mathbf{Q}_p)$.) We remark that the introduction of adeles into the study of algebraic groups is due to Weil [Wei82].

Definition 4.7.1. An *automorphic form on* GL_n is a function

$$\phi\colon \mathrm{GL}_n(\mathbf{Q})\backslash\mathrm{GL}_n(\mathbf{A})\to\mathbf{C}$$

which is square-integrable relative to the pushforward of a Haar measure on $\mathrm{GL}_n(\mathbf{A})$ to the quotient $\mathrm{GL}_n(\mathbf{Q})\backslash\mathrm{GL}_n(\mathbf{A})$. Let L^2 be the space of automorphic forms on GL_n. An *automorphic representation of* GL_n is an irreducible representation of $\mathrm{GL}_n(\mathbf{A})$ appearing in L^2.

The notion of "appearing in" may be rather subtle owing to the existence of the continuous series, as in the example above with $X = G = \mathbf{R}$. Also there is a notion of *cuspidal* automorphic form, involving a vanishing of certain zeroth Fourier coefficients; an automorphic representation is called cuspidal if it appears in the space L^2_0 of cuspidal automorphic forms. Finally, there is a more general notion of automorphic form on G, where G is any *reductive algebraic group*.

For formal reasons, an irreducible representation π of $\mathrm{GL}_n(\mathbf{A})$ decomposes as a "restricted tensor product" $\bigotimes'_{p\leq\infty}\pi_\infty$, where π_p is an irreducible representation of $\mathrm{GL}_n(\mathbf{Q}_p)$. Furthermore, for almost all p, π_p is what is known as an *unramified principal series representation*. For $p\neq\infty$, the unramified principal series representations of $\mathrm{GL}_n(\mathbf{Q}_p)$ are parametrized by unordered n-tuples of complex numbers $\{\alpha_{p,1},\dots,\alpha_{p,n}\}$ (the *Satake parameters*). For each local π_p there is an L-factor $L(\pi_p,s)$, which for unramified principal series representations is

$$L(\pi_p,s)=\prod_{i=1}^{n}(1-\alpha_{p,i}p^{-s})^{-1}.$$

If π is a cuspidal automorphic representation, the global L-function $L(\pi,s)=\prod_p L_p(\pi,s)$, convergent for $\mathrm{Re}\ s\gg 0$, admits an analytic continuation and functional equation [GJ72].

Example 4.7.2 (Dirichlet characters as automorphic forms, and Hecke characters)**.** If $n=1$, an automorphic form is a function on $\mathrm{GL}_1(\mathbf{Q})\backslash\mathrm{GL}_1(\mathbf{A})=\mathbf{Q}^\times\backslash\mathbf{A}^\times$. We claim that

$$\mathbf{Q}^\times\backslash\mathbf{A}^\times=\prod_{p<\infty}\mathbf{Z}_p^\times\times\mathbf{R}^\times_{>0}.$$

Indeed, if $a=(a_p)_{p\leq\infty}\in\mathbf{A}^\times$, we can find a unique rational number γ so that $|\gamma|_p=|a_p|_p$ for all $p<\infty$ and $\gamma a_\infty>0$. (Explicitly, $\gamma=\mathrm{sgn}(a_\infty)\prod_{p<\infty}|a_p|_p^{-1}$.) Then $\gamma^{-1}a\in\prod_{p<\infty}\mathbf{Z}_p^\times\times\mathbf{R}^\times_{>0}$. An automorphic representation of GL_1 is called a *Hecke character*. It is just a character $\chi\colon\mathbf{Q}^\times\backslash\mathbf{A}^\times\cong\prod_{p<\infty}\mathbf{Z}_p^\times\times\mathbf{R}^\times_{>0}\to\mathbf{C}^\times$.

Hecke characters may have infinite order. If χ has finite order, it vanishes on $\mathbf{R}^\times_{>0}$, and must therefore factor through some finite quotient group of $\prod_{p<\infty}\mathbf{Z}_p^\times$, which by the Chinese remainder theorem is isomorphic to $(\mathbf{Z}/N\mathbf{Z})^\times$ for some $N\geq 1$. Thus *finite-order Hecke characters are essentially the same as Dirichlet characters*. Every Hecke character takes the form $\chi_0\chi_\infty$, where χ_0 is finite order and χ_∞ factors through a continuous homomorphism $\mathbf{R}^\times_{>0}\to\mathbf{C}^\times$.

These notions extend to a general number field K. Recall that a place of K is either a prime of K or an embedding of K into $\mathbf{C}$. For each place $\mathfrak{p}$ of K we have

the completion $K_{\mathfrak{p}}$. When $\mathfrak{p}$ is infinite, $K_{\mathfrak{p}}$ is $\mathbf{R}$ or $\mathbf{C}$. When $\mathfrak{p}$ is finite, $K_{\mathfrak{p}}$ is a finite extension of $\mathbf{Q}_p$ for some prime number p. Let $\mathcal{O}_{\mathfrak{p}}$ be its ring of integers, and let $\varpi \in \mathcal{O}_{\mathfrak{p}}$ be a prime element. If χ is a Hecke character of K, its L-function is

$$L(\chi, s) = \prod_{\mathfrak{p} \text{ unram.}} \left(1 - \frac{\chi(\varpi_{\mathfrak{p}})}{N\mathfrak{p}^s}\right)^{-1}.$$

The analytic continuation and functional equation of $L(\chi, s)$ was established by Hecke himself, but Tate's thesis [Cas67, Ch. XV] gave a new proof using harmonic analysis on adele groups. Tate's thesis laid the foundations for the modern theory of automorphic representations.

Example 4.7.3 (Modular forms as automorphic forms)**.** Let g be a modular form of weight k. We may define an automorphic form ϕ_g on GL_2 through a formula similar to (4.7.1), see [Gel75, §3]. If g is a cuspidal eigenform, the representation of $\mathrm{GL}_2(\mathbf{A})$ spanned by translates of ϕ_g is an automorphic representation π_g. One gets a correspondence $g \mapsto \pi_g$ between the set of cuspidal eigenforms and the set of cuspidal automorphic representations of GL_2 which are discrete series at ∞. Under mild conditions on g, one has an equality of L-functions $L(g, s) = L(\pi_g, s)$.

Conjecture 4.7.4 (Langlands reciprocity)**.** *Let π be a cuspidal algebraic*[8] *automorphic representation of* GL_n*. Let p be a prime, and let $\iota\colon \mathbf{C} \to \overline{\mathbf{Q}}_p$ be a field isomorphism. Then there exists an irreducible p-adic Galois representation*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p),$$

such that for almost all primes ℓ, ρ is unramified at ℓ and the roots of the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ are the images under ι of the Satake parameters $\alpha_{\ell,1}, \ldots, \alpha_{\ell,n}$ of π_ℓ. Furthermore, ρ is "geometric".

The converse to Conjecture 4.7.4 is the Fontaine–Mazur conjecture, which generalizes Conjecture 4.5.2. Thus there is a conjectural bijection between algebraic cuspidal automorphic representations of GL_n and a certain class of n-dimensional p-adic representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Both conjectures remain out of reach in the level of generality we have posed them. The situation is much better understood when $\mathbf{Q}$ is replaced with the function field of a curve defined over a finite field. There the analogues of Conjecture 4.7.4 and its converse were established by L. Lafforgue [Laf02], building on work of Drinfeld.

Example 4.7.5 (Algebraic Maass forms)**.** The simplest case of Conjecture 4.7.4 not accessible by current methods occurs when π is a cuspidal algebraic automorphic representation of GL_2 for which π_∞ is a so-called principal series representation. Such π correspond to algebraic Maass forms. A *Maass form* is an analytic (not holomorphic) function on $\Gamma_1(N)\backslash\mathcal{H}$ which is an eigenvector for the Laplacian operator $y^{-2}(\partial^2/\partial x^2 + \partial^2/\partial y^2)$; it is *algebraic* if the eigenvalue is $1/4$. A finer form of Conjecture 4.7.4 (see [Gel97, §2]) predicts a correspondence between two-dimensional even Artin representations and cuspidal algebraic Maass eigenforms. Nobody has any idea how to prove the correspondence in those instances where the Artin representation is of icosahedral type. As an example, the polynomial

$$f(x) = x^5 - x^4 - 780x^3 - 1795x^2 + 3106x + 344$$

[8] "Algebraic" is a certain condition on π_∞ that we will not define here; in the case that $n = 1$ and $\pi = \chi = \chi_0\chi_\infty$ is a Hecke character, it is the condition that $\chi_\infty(z) = z^k$ for an integer k. See [Clo90] for the precise definition.

has Galois group A_5. As in Example 4.4.2, there is a corresponding Artin representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$; it happens to be unramified outside the single prime 1951 (see [DM06]). But this time $f(x)$ has all real roots, so ρ is even. Conjecturally, there exists a Maass form

$$g(x+iy) = \sum_{n\neq 0} a_n \sqrt{y} K_0(2\pi\,|n|\,y) e^{2\pi i n x}$$

for the group $\Gamma_1(1951)$ such that $a_p = \operatorname{tr}\rho(\mathrm{Frob}_p)$ for all primes $p \neq 1951$. Here $K_0(y)$ is a Bessel function:

$$K_0(y) = \frac{1}{2}\int_0^\infty \exp\left(-\frac{y}{2}\left(t+\frac{1}{t}\right)\right)\frac{dt}{t},\ y>0.$$

5. The cohomology of arithmetic manifolds

5.1. **Arithmetic manifolds for GL_n.** In §4.6 we discussed the special role that modular curves play in the proof of Theorem 4.4.1. Recall that the Riemann surface $Y_1(N)$ is the quotient $\Gamma_1(N)\backslash\mathcal{H}$, where $\Gamma_1(N) \subset \mathrm{SL}_2(\mathbf{Z})$ is the finite-index subgroup defined in (4.1.3) and $X_1(N)$ is a compactification of $Y_1(N)$. For every prime $p \nmid N$, the Hecke operators T_p and $\langle p\rangle$ act by algebraic correspondences on $X_1(N)$. Therefore they act as a commuting family of endomorphisms on the singular cohomology group $H^1(X_1(N), \mathbf{Q})$, and it makes sense to talk about a *Hecke eigenclass* in this space, possibly after extending scalars to a finite extension of $\mathbf{Q}$. The proof of Theorem 4.4.1 (for modular forms of weight 2, anyway) involved a combination of two facts:

1. Hecke eigenclasses in $H^1(X_1(N), \mathbf{Q})$ correspond to cuspidal eigenforms.
2. The p-adic cohomology $H^1(X_1(N), \mathbf{Q}_p)$ can be interpreted as an étale cohomology group of an algebraic curve over $\mathbf{Q}$, and therefore it admits an action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which commutes with the action of the Hecke operators.

One might seek to generalize Theorem 4.4.1 to higher dimensions as follows. The group $\mathrm{SL}_2(\mathbf{R})$ acts transitively on $\mathcal{H}$, and the stabilizer of i is $\mathrm{SO}(2)$, so $\mathcal{H} \cong \mathrm{SL}_2(\mathbf{R})/\mathrm{SO}(2)$. Let us put $\mathcal{H}_n = \mathrm{SL}_n(\mathbf{R})/\mathrm{SO}(n)$; this is a manifold with a left action by $\mathrm{SL}_n(\mathbf{R})$. One can form the quotient $\Gamma\backslash\mathcal{H}_n$ by a congruence subgroup $\Gamma \subset \mathrm{SL}_n(\mathbf{Z})$ whose definition is analogous to $\Gamma_1(N)$. The result is an example of an *arithmetic manifold*. As with modular curves, arithmetic manifolds admit Hecke correspondences. For each prime $p \nmid N$, there are n Hecke correspondences $T_{p,1}, \ldots, T_{p,n}$. Let us only say that when $n = 2$, $T_{p,1}$ and $T_{p,2}$ are T_p and $\langle p\rangle$, respectively. For each $j \geq 0$, the Hecke correspondences act as endomorphisms on $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Q})$, and so one can talk about Hecke eigenclasses. Do these correspond to n-dimensional Galois representations?

Fact (1) above generalizes nicely to our situation: Hecke eigenclasses in the cohomology of $\Gamma\backslash\mathcal{H}_n$ (possibly with coefficients in a nontrivial local system arising from an algebraic representation of GL_n) correspond to automorphic representations of GL_n. The correspondence only sees automorphic representations of a certain sort known as *cohomological*.[9] We will only say here that the condition that π be cohomological is a condition on the infinite component π_∞, and that the precise relationship between eigenclasses in $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Q})$ and cohomological representations is known (we are referring to *Matsushima's formula*, see [BW00, Ch. VII]). If

[9] Also known as *regular algebraic*.

we were able to associate a Galois representation to an eigenclass in $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Q})$, it would establish Langlands reciprocity (Conjecture 4.7.4) for the corresponding automorphic representation.

Generalizing (2), we "hit a wall" immediately for $n > 2$, however. The problem is that $\mathcal{H}_n$ is not a complex manifold for $n > 2$, and so no quotient of it is going to be an algebraic variety. For instance, $\mathcal{H}_3$ has dimension 5, which is *odd*. Finding a Galois representation seems rather hopeless. Nonetheless, the following theorem was announced around 2012:

Theorem 5.1.1 ([HLTT], [Sch13c])**.** *Let g be a Hecke eigenclass in the singular cohomology $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{C})$, and let $a_{\ell,i}(g)$ be the eigenvalue of $T_{\ell,i}$ on g for $\ell \nmid N$ prime and $i = 1, \dots, n$. There exists a continuous semisimple p-adic Galois representation*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p)$$

which is associated to g in the sense that for all primes $\ell \nmid Np$, ρ is unramified at ℓ, and the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is

$$x^n + \sum_{k=1}^{n} (-1)^k \ell^{k(k-1)/2} a_{\ell,k}(g) x^{n-k}. \tag{5.1.1}$$

The results of [HLTT] and [Sch13c] are rather stronger than this: they show that every cuspidal regular algebraic automorphic representation of GL_n over a totally real or CM field F has an associated Galois representation. Theorem 5.1.1 is the special case $F = \mathbf{Q}$.

5.2. **Scholze's theorem on torsion classes.** In fact the results of [Sch13c] are stronger still. Theorem 5.1.1 concerns the singular cohomology $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{C})$ with complex coefficients, but we could also have considered the integral cohomology $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Z})$, a finitely generated abelian group equipped with the action of Hecke operators $T_{p,i}$. When $n = 2$, $Y_1(N) = \Gamma_1(N)\backslash\mathcal{H}$ is a surface; the integral cohomology groups of a surface are known to be torsion-free. But for $n > 2$, the cohomology $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Z})$ can contain a large torsion subgroup. This torsion subgroup is also preserved by the Hecke operators. Ash [Ash92, Conjecture B] asked whether the *mod p* eigenclasses have corresponding *mod p* Galois representations. In fact they do:

Theorem 5.2.1 ([Sch13c])**.** *Let p be prime, and let g be a Hecke eigenclass in $H^j(\Gamma\backslash\mathcal{H}_n, \mathbf{Z}/p\mathbf{Z})$. Then there exists a continuous semisimple Galois representation*

$$\rho_g\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{F}}_p)$$

which is associated to g in the same sense as in Theorem 5.1.1*, except that the polynomial in* (5.1.1) *now has coefficients in $\overline{\mathbf{F}}_p$.*

Theorem 5.2.1 is a partial answer to Question C for the Galois representations ρ_g. We remark that Theorem 5.2.1 also applies to eigenclasses in $H^j(\Gamma\backslash\mathcal{H}_n, V)$, where V is a local system. In prior years, Ash and others had developed a conjectural converse to Theorem 5.2.1, which predicts that every Galois representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{F}}_p)$ satisfying a "strict parity condition" (which generalizes the notion of being odd) has a corresponding Hecke eigenclass g in $H^j(\Gamma\backslash\mathcal{H}_n, V)$ for an appropriate choice of Γ, j, and V. See for instance [ADP02], which offers a great deal of numerical evidence.

5.3. **Arithmetic manifolds in the large.** The rest of the article will be an exposition of the ideas behind [Sch13c]. We begin with a discussion of arithmetic manifolds, of which $\Gamma\backslash\mathcal{H}_n$ is an example.

An *arithmetic manifold* is a double coset space

$$X = \Gamma\backslash G(\mathbf{R})/K,$$

where

- G is a semisimple algebraic group over $\mathbf{Q}$,
- K is a maximal compact subgroup of the Lie group $G(\mathbf{R})$, and
- $\Gamma \subset G(\mathbf{R})$ is an *arithmetic subgroup*—this is a generalization of the notion of a finite-index subgroup of $\mathrm{SL}_2(\mathbf{Z})$; see [Mil13, Ch. VII] for a precise definition.

There is a natural way to give $D = G(\mathbf{R})/K$ the structure of a Riemannian manifold. One shows that D is a *Riemannian symmetric space*: this means that for every point $p \in D$ there exists an isometry $i_p \colon D \to D$ which fixes p and whose derivative at p is multiplication by -1.

One can also show that an arithmetic subgroup $\Gamma \subset G(\mathbf{R})$ is a *lattice* (meaning a discrete subgroup with finite covolume with respect to the Haar measure on $G(\mathbf{R})$). The quotient $X = \Gamma\backslash D$ is a *locally Riemannian symmetric space*, meaning that for every point $p \in X$ there exists an isometry i_p as in the previous paragraph, except that it may only be defined in a neighborhood of p. Furthermore, X has finite volume.

Conversely, suppose that X is a locally Riemannian symmetric space of finite volume. Let $\widetilde{X}$ be its universal cover. One can show that the identity component of $\mathrm{Aut}\,\widetilde{X}$ is a semisimple real Lie group $\mathcal{G}$, and that $\widetilde{X} \cong \mathcal{G}/K$ for $K \subset \mathcal{G}$ a maximal compact subgroup. Thus $X = \Gamma\backslash\mathcal{G}/K$ for a *lattice* $\Gamma \subset \mathcal{G}$. Let us call a lattice $\Gamma \subset \mathcal{G}$ "arithmetic" if X is an arithmetic manifold. The following is an incredible theorem of Margulis.

Theorem 5.3.1 ([Mar91])**.** *As long as $\mathcal{G}$ has no factor isogenous to* $\mathrm{SO}(\mathrm{n},1)$ *or* $\mathrm{SU}(\mathrm{n},1)$*, every lattice in $\mathcal{G}$ is arithmetic.*

Thus outside of the exceptional cases described by the theorem, every locally Riemannian symmetric space of finite volume is an arithmetic manifold! In particular every lattice in $\mathrm{SL}_n(\mathbf{R})$ for $n \geq 3$ is arithmetic. In contrast $\mathrm{SL}_2(\mathbf{R})$ and $\mathrm{SL}_2(\mathbf{C})$ (which are isogenous to $\mathrm{SO}(1,1)$ and $\mathrm{SO}(3,1)$, respectively) have uncountably many conjugacy classes of nonarithmetic subgroups.

The arithmetic manifolds relevant to reciprocity laws are those where Γ is a *congruence subgroup*, meaning that Γ contains a subgroup of the form $\ker(G(\mathbf{Z}) \to G(\mathbf{Z}/N\mathbf{Z}))$ for some integer N. There are infinitely many conjugacy classes of finite-index subgroups of $\mathrm{SL}_2(\mathbf{Z})$ which are not congruence subgroups. But once again SL_2 is exceptional. A theorem proved independently by Bass–Lazard–Serre and Mennicke shows that for $n \geq 3$, every finite-index subgroup of $\mathrm{SL}_n(\mathbf{Z})$ is a congruence subgroup. For a discussion of the "congruence subgroup problem", see [Rag04].

If Γ is a congruence subgroup, then the arithmetic manifold $\Gamma\backslash G(\mathbf{R})/K$ has Hecke correspondences for almost every prime p, and so there is the possibility of posing a version of Theorem 5.1.1 for Hecke eigenclasses in the cohomology of this space. There is an adelic construction of these manifolds which is more in line with

the philosophy of automorphic representations. Let G be a reductive group, and let $X = G(\mathbf{Q})\backslash G(\mathbf{A})/K_0K_\infty$, where $K_0 \subset G(\mathbf{A}_f)$ is a compact open subgroup and $K_\infty \subset G(\mathbf{R})$ is a maximal compact subgroup. Then Matsushima's formula [BW00] relates the cohomology of X to cohomological automorphic representations of G.

Example 5.3.2 (Bianchi manifolds)**.** Let $\mathcal{G} = \mathrm{SL}_2(\mathbf{C})$. Then $K = \mathrm{SU}(2)$ is a maximal compact subgroup of $\mathcal{G}$, and $\mathcal{G}/K \cong \mathbf{C} \times \mathbf{R}_{>0}$ is hyperbolic 3-space. Let F be an imaginary quadratic field with ring of integers $\mathcal{O}_F$ (e.g., $F = \mathbf{Q}(i)$ and $\mathcal{O}_F = \mathbf{Z}[i]$). If Γ is a congruence subgroup of $\mathrm{SL}_2(\mathcal{O}_F)$, then $X = \Gamma\backslash\mathcal{G}(\mathbf{R})/K$ is called a *Bianchi manifold*. (It is indeed an arithmetic manifold; the role of G is played by the Weil restriction of SL_2/F from F to $\mathbf{Q}$.) In fact Theorem 5.2.1 generalizes to such X: Hecke eigenclasses in the mod p cohomology of X correspond to two-dimensional mod p representations of $\mathrm{Gal}(\overline{F}/F)$. See [Sen14] for a survey of the arithmetic of Bianchi manifolds.

5.4. **Shimura varieties.** We are especially interested in arithmetic manifolds $\Gamma\backslash G(\mathbf{R})/K$, which are algebraic varieties defined over a number field, in the hopes of constructing Galois representations. A necessary condition for this is that the Riemannian symmetric space $G(\mathbf{R})/K$ must have a complex structure compatible with its Riemannian structure; i.e., it must be a *Hermitian symmetric domain*. This occurs if and only if K contains a central subgroup isomorphic to the circle group S^1; a quarter turn by this circle furnishes the complex structure on the identity coset of $G(\mathbf{R})/K$. Examples of G for which $G(\mathbf{R})/K$ is a Hermitian symmetric domain include Sp_{2n}, $\mathrm{U}(p,q)$ and $\mathrm{O}(2,n)$. Note that $\mathrm{Sp}_2 = \mathrm{SL}_2$, and that the corresponding Hermitian symmetric domain is the upper half-plane $\mathcal{H}$.

The following "meta-theorem" encompasses a series of important results by Shimura, which were put into a common perspective by Deligne. It generalizes the fact that $\Gamma_1(N)\backslash\mathcal{H}$ is an algebraic curve defined over $\mathbf{Q}$.

Theorem 5.4.1 ([Del79])**.** *In many cases, the quotient of a Hermitian symmetric domain by a congruence subgroup is an algebraic variety defined over a number field, which can be given explicitly.*

Varieties constructed this way are called *Shimura varieties*, and they provide the vital link between automorphic representations and Galois representations. Because of Theorem 5.4.1, there is hope of replicating some of the theory of elliptic modular forms discussed in §4 in the context of Shimura varieties.

Example 5.4.2 (Siegel modular varieties)**.** Let $\Gamma \subset \mathrm{Sp}_{2n}(\mathbf{Z})$ be a congruence subgroup. The arithmetic manifold

$$\mathrm{Sh}_\Gamma = \Gamma\backslash \mathrm{Sp}_{2n}(\mathbf{R})/\mathrm{U}(n)$$

is a Shimura variety known as a *Siegel modular variety*. It is the moduli space for principally polarized abelian varieties of dimension n equipped with a level structure dictated by Γ. A *Siegel cusp form* g of weight k and level Γ is a holomorphic function on the complex manifold $\mathrm{Sp}_{2n}(\mathbf{R})/\mathrm{U}(n)$ which transforms appropriately under Γ, and which vanishes at the cusps of Sh_Γ in an appropriate sense. Alternatively g can be seen as a section of the line bundle $\omega^{\otimes k}$, where ω is the push-forward of the canonical line bundle through the universal abelian variety $\mathcal{A} \to \mathrm{Sh}_\Gamma$. For almost every prime ℓ, there are Hecke operators $T_{\ell,1}, \ldots, T_{\ell,2n+1}$ acting on the space of Siegel modular forms of weight k and level Γ.

Theorem 5.4.3. *Let g be a Siegel cusp form which is an eigenform for all the Hecke operators. Then for each prime p there is a Galois representation*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_{2n+1}(\overline{\mathbf{Q}}_p)$$

which is associated to g in the sense that for almost all primes ℓ, the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is determined by the eigenvalues of the operators $T_{\ell,i}$ on g.

We refer to [Sch13c, Cor. V.1.7] for the precise statement. The proof of Theorem 5.4.3, which combines contributions from many people, is far more complicated than that of Theorem 4.4.1. We will only say that after applying heavy automorphic machinery of Arthur, the Galois representation ρ in this theorem is found within the cohomology of an appropriate Shimura variety.

The "many cases" of arithmetic manifolds referred to by Theorem 5.4.1 are essentially those which can be embedded inside of the Siegel modular variety. They include the cases where G is $\mathrm{U}(p,q)$ and $\mathrm{O}(2,n)$.

5.5. **Arithmetic manifolds at the boundary of a Shimura variety.** Let Sh be the Shimura variety of Example 5.4.2, corresponding to a congruence subgroup $\Gamma \subset \mathrm{Sp}_{2n}(\mathbf{Z})$, and let X be an arithmetic manifold for GL_n, corresponding to a congruence subgroup $\Gamma_0 \subset \mathrm{SL}_n(\mathbf{Z})$. The proof of Theorem 5.2.1 leverages a topological connection between Sh (which is an algebraic variety) and X (which is not, in general). The connection comes from the fact that Sp_{2n} contains a parabolic subgroup P, consisting of matrices of the form

$$\begin{pmatrix} A & B \\ 0 & {}^tA^{-1} \end{pmatrix},\ A \in \mathrm{GL}_n, B \in M_{n\times n}, A^tB \text{ symmetric.}$$

The Levi subgroup of P is isomorphic to GL_n. Let us now assume that the projection $P \to \mathrm{GL}_n$ carries $\Gamma \cap P$ onto Γ_0.

The Shimura variety Sh is not compact. There are many ways to compactify it; for our purposes we need the compactification $\overline{\mathrm{Sh}}$ due to Borel and Serre ([BS73], see also the exposition in [Gor05, §4], and the book [BJ06]). This compactification has the following properties:

- $\overline{\mathrm{Sh}}$ is a compact manifold with corners;
- The inclusion $\mathrm{Sh} \hookrightarrow \overline{\mathrm{Sh}}$ is a homotopy equivalence;
- The boundary $\overline{\mathrm{Sh}} \setminus \mathrm{Sh}$ is a stratified manifold, with each stratum a locally symmetric space for a parabolic subgroup of Sp_{2n};
- In particular the boundary contains X_P as an open subset, where X_P is a torus bundle over X.

The result is that cohomology classes on X appear in Sh. The compactification $\overline{\mathrm{Sh}}$ is preserved by Hecke correspondences, so that if g is a Hecke eigenclass in $H^i(X, \mathbf{Z}/p\mathbf{Z})$, then there exists an eigenclass g' in $H^i(\mathrm{Sh}, \mathbf{Z}/p\mathbf{Z})$ whose Hecke eigenvalues are related to those of g in a systematic way. The precise statement, which is rather technical, is [Sch13c, Cor. V.2.4].

We wish to produce a Galois representation associated to g'. It would be nice if g' were the image of a torsion-free eigenclass in $H^i(\mathrm{Sh}, \mathbf{Z})$, for then there would be a corresponding automorphic representation by Matsushima's formula, and then we could apply Theorem 4.4. However this may not be the case, as $H^i(\mathrm{Sh}, \mathbf{Z})$ will typically have lots of torsion elements.

Scholze's breakthrough comes in the form of the following theorem.

Theorem 5.5.1 ([Sch13c, Theorem I.5])**.** *Let g' be an eigenclass in the cohomology group $H^i(\mathrm{Sh}, \mathbf{Z}/p\mathbf{Z})$. Then there exists a Siegel cusp form h (possibly for a subgroup smaller than Γ) whose Hecke eigenvalues are congruent to those of g' modulo p.*

Granting Theorem 5.5.1, we can now indicate how the proof of Theorem 5.2.1 proceeds. Given an eigenclass $g \in H^i(X, \mathbf{Z}/p\mathbf{Z})$, there exists a corresponding eigenclass $g' \in H^i(\mathrm{Sh}, \mathbf{Z}/p\mathbf{Z})$, and then by Theorem 5.5.1 there is a corresponding cusp form h. Applying Theorem 5.4.3 one finds a p-adic Galois representation ρ' of dimension $2n+1$. The reduction $\rho' \pmod p$ is related to the Hecke eigenvalues on g. However it is $(2n+1)$ dimensional, not n dimensional. The final arguments of [Sch13c] show that the origins of g as a cohomology class on X (rather than on Sh) place constraints on $\rho' \pmod p$, forcing the existence of an n-dimensional summand ρ of $\rho' \pmod p$. This ρ is the mod p Galois representation required by Theorem 5.2.1.

5.6. **Rigid-analytic spaces and their cohomology.** Theorem 5.5.1 asserts a connection between a *topological object* (a cohomology class) and an *analytic object* (a cusp form) associated with the Siegel modular variety Sh. Such connections are well known in the world of classical manifolds. The most basic example is the de Rham theorem,

$$H^k(X, \mathbf{Z}) \otimes_{\mathbf{Z}} \mathbf{R} \cong H^k_{\mathrm{dR}}(X),$$

which connects the topology of a compact manifold X with differential forms on it. If X is a compact Kähler manifold (which is the case if X is a projective variety), then there is a Hodge decomposition

$$H^k(X, \mathbf{Z}) \otimes_{\mathbf{Z}} \mathbf{C} \cong \bigoplus_{i+j=k} H^i(X, \Omega^j_{X/\mathbf{C}}), \tag{5.6.1}$$

where $\Omega^j_{X/\mathbf{C}}$ is the sheaf of holomorphic j-forms on X. However, the Hodge decomposition of the Shimura variety Sh will be of no use in proving Theorem 5.5.1, since the desired connection between the topological object and the analytic object is a *congruence*, which makes no sense in the context of complex vector spaces.

Instead we turn to the theory of *rigid-analytic spaces*, which runs parallel to the theory of complex manifolds, but in which the field $\mathbf{C}$ is replaced by a p-adic field K. The theory was first developed by Tate [Tat71]; a standard reference is [BGR84]. In this theory one defines a *K-affinoid algebra* A, which is a certain kind of Banach K-algebra. The set of maximal ideals $\mathrm{Spm}\, A$ is called an *affinoid space*, and a general rigid-analytic space is created by gluing together affinoid spaces, as schemes are created by gluing together affine schemes.

Example 5.6.1 (The rigid-analytic closed disc)**.** Let C be the completion of an algebraic closure of $\mathbf{Q}_p$, and let $A = C\langle T \rangle$ be the ring of power series $f = \sum_{n \geq 0} a_n T^n$ with C coefficients such that $a_n \to 0$. The unusual topology of p-adic numbers implies that a power series with C coefficients converges on the closed unit disc $D = \{z \in C \,|\, |z| \leq 1\}$ if and only if it belongs to A. In fact $D \cong \mathrm{Spm}\, A$ via $z \mapsto \ker(f \mapsto f(z))$.

Let $K/\mathbf{Q}_p$ be a topological field which is complete with respect to an absolute value $z \mapsto |z|$ extending the one on $\mathbf{Q}_p$. A K-affinoid algebra A is defined to be any quotient of $K\langle T_1, \ldots, T_n \rangle$. This ring (called a *Tate algebra*) is noetherian. Therefore a general affinoid space can be visualized as a closed subset of a polydisc

$\{(z_1, \ldots, z_n) | \, |z_i| \leq 1\}$ cut out by a finite list of equations given by convergent power series.

Does a version of (5.6.1) hold for rigid-analytic varieties X? This question was posed by Tate himself [Tat67], who answered it affirmatively in the case that X is an abelian variety. To even pose the question, one needs an analogue of $H^i(X, \mathbf{Z})$ for a rigid-analytic space. Singular cohomology will not behave well for rigid-analytic spaces, because their topology is not anything like a classical manifold. As with schemes, rigid-analytic spaces have an étale site [Hub96], so that one can define cohomology groups $H^i(X_{\text{ét}}, \mathbf{Z}/n\mathbf{Z})$ for any integer n, as well as the p-adic cohomology $H^i(X_{\text{ét}}, \mathbf{Z}_p)$. If X starts out life as a nonsingular variety over $\mathbf{Q}$, then there is a corresponding rigid-analytic space X^{an}, and the various cohomology theories attached to X agree:

$$H^i(X(\mathbf{C}), \mathbf{Z}) \otimes_{\mathbf{Z}} \mathbf{Z}_p \cong H^i(X_{\overline{\mathbf{Q}},\text{ét}}, \mathbf{Z}_p) \cong H^i(X^{\text{an}}_{C,\text{ét}}, \mathbf{Z}_p). \tag{5.6.2}$$

Tate's question in [Tat67] asks whether the p-adic étale cohomology groups of a proper rigid-analytic variety satisfy an analogue of the Hodge decomposition. It was answered affirmatively for proper algebraic varieties by Fontaine and Messing [FM87] and Faltings [Fal88] and in general by Scholze [Sch13a]:

Theorem 5.6.2. *Let X be a rigid-analytic space over $\mathbf{Q}_p$ which is smooth and proper (a condition akin to being compact). Then X has a "Hodge–Tate decomposition"*

$$H^k(X_{C,\text{ét}}, \mathbf{Z}_p) \otimes_{\mathbf{Z}_p} C \cong \bigoplus_{i+j=k} H^i(X, \Omega^j_{X/\mathbf{Q}_p}) \otimes_{\mathbf{Q}_p} C(-j).$$

The Hodge decomposition is compatible with the action of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. The $C(-j)$ refers to a one-dimensional C vector space on which $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ acts through the $-j$th power of the p-adic cyclotomic character. Scholze also shows that the action of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ on $H^k(X_{C,\text{ét}}, \mathbf{Q}_p)$ is "geometric" in the same sense as in Conjecture 4.5.2.

5.7. **Some remarks on p-adic geometry, and the inevitability of perfectoid spaces.** A fundamental difference between classical manifolds and rigid-analytic spaces is that the former are locally contractible: A manifold X can be covered by open subsets U_i such that all intersections between the U_i are contractible. Consequently if $\mathcal{F}$ is a sheaf on X, computing the cohomology groups $H^i(X, \mathcal{F})$ can be reduced to a combinatorial study of the values of $\mathcal{F}$ on the intersections of the U_j. In other words, *the cohomology of a manifold can be computed by the Čech complex of a sufficiently fine open covering.*

For a general (not necessarily contractible) open cover $\mathcal{U} = \{U_i\}$, the Čech complex of $\mathcal{F}$ with respect to $\mathcal{U}$ does not compute $H^i(X, \mathcal{F})$; the failure to do so is measured by the cohomology of $\mathcal{F}$ on the intersections of the U_i. The precise statement is that we have a spectral sequence

$$\check{H}^i(\mathcal{U}, \mathcal{H}^j(\mathcal{F})) \implies H^{i+j}(X, \mathcal{F}),$$

where $\mathcal{H}^j(\mathcal{F})$ is the presheaf $U \mapsto H^j(U, \mathcal{F})$. If $H^j(U_I, \mathcal{F}) = 0$ for all intersections U_I among the U_i and all $j > 0$, then the spectral sequence simply gives $\check{H}^i(\mathcal{U}, \mathcal{F}) = H^i(X, \mathcal{F})$.

The same formalism applies to sheaves on any topological space or site[10] X. Suppose $\mathcal{F}$ is a sheaf on X. If one can find a covering $\mathcal{U} = \{U_i\}_{i \in I}$ of X such that $H^i(U_J, \mathcal{F}) = 0$ for all finite subsets $J \subset I$ and all $i > 1$, then $\check{H}^i(\mathcal{U}, \mathcal{F}) = H^i(X, \mathcal{F})$.

Example 5.7.1 (Quasi-coherent sheaves)**.** Let X be a scheme, and let $\mathcal{F}$ be a quasi-coherent sheaf on U. For example $\mathcal{F}$ could be the structure sheaf $\mathcal{O}_X$. If X is affine, then $H^j(X, \mathcal{F}) = 0$ for all $j > 0$. (In fact the converse is true: this is Serre's criterion for affineness.) If X is an arbitrary separated scheme, then one can find an open cover $\mathcal{U} = \{U_i\}$ of X such that all finite intersections among the U_i are affine. Then the cohomology of $\mathcal{F}$ can be computed by the Čech complex of $\mathcal{U}$.

If X is a rigid-analytic space, and $\mathcal{F}$ is a sheaf on the étale site of X, one might hope that there exists an étale covering $\{U_i \to X\}$ which is fine enough so that $\mathcal{F}$ has no cohomology on the fiber products among the U_i. This is false for the constant sheaf $\mathbf{Z}/p\mathbf{Z}$, as the following example suggests.

Example 5.7.2 (The rigid-analytic closed disc is not contractible)**.** Let $D = \operatorname{Spm} C\langle T \rangle$ be the closed disc from Example 5.6.1. Let $f(T) \in C\langle T \rangle$ have coefficients bounded by 1, so that $|f(z)| \leq 1$ for all $z \in D$. The equation $Y^p - Y = f$ defines an étale cover $D' \to D$ (an *Artin–Schreier* cover), because its Y-derivative $pY^{p-1} - 1$ is nowhere zero on D'. For many values of f this cover is connected, which indicates that the étale fundamental group of D is quite large. So in the p-adic world, the closed disc isn't simply connected! In particular $H^1(D_{\text{ét}}, \mathbf{Z}/p\mathbf{Z})$ is not even finitely generated. The same should be true for all affinoid spaces of positive dimension.

Thus the open sets $U \to X$ constituting the étale topology are not fine enough to compute the cohomology of a rigid-analytic space. In [Sch13a], Scholze sidesteps this problem by defining a finer topology on X, the *pro-étale site*, in which opens $U \to X$ are allowed to be *infinite-to-one*. They are defined as $U = \varprojlim U_n$, where each $U_n \to X$ is étale. But then what sort of a beast is U? The following example is not to be taken too seriously; rather the intent is to leave an impression on the reader of how strange such a U can be.

Example 5.7.3 (The p-adic solenoid)**.** Let $X = S^1$, and for each $n \geq 0$ let $X_n \to X$ be the p^n-fold cover $x \mapsto x^{p^n}$. Then the X_n form a projective system, and one can define the space $\widetilde{X} = \varprojlim X_n$ with the inverse limit topology. Thus X is the set of sequences $(z_0, z_1, \dots)$ of complex numbers with $|z_i| = 1$ and $z_i^p = z_{i-1}$ for all $i \geq 1$. (In fact X is isomorphic to the quotient of $\mathbf{R} \times \mathbf{Z}_p$ modulo a diagonally embedded $\mathbf{Z}$.) Then $\widetilde{X}$ is not a manifold at all. For instance it is not path-connected (although it is connected).

Example 5.7.4 (The perfectoid closed disc)**.** Let $X = D$ as in Example 5.6.1, and let $X_n \to X$ be the map $D \to D$ given by $x \mapsto x^{p^n}$. This map is étale away from the origin. Let $\widetilde{D} = \varprojlim X_n$; then $\widetilde{D} \setminus \{0\}$ is a pro-étale cover of $D \setminus \{0\}$. However, $\widetilde{D}$ does not exist in the category of rigid-analytic spaces.

Extracting arbitrary pth power roots of functions is a common technique in p-adic geometry, e.g., [Fal02]. But it is to Scholze's credit that he could fearlessly incorporate strange spaces like $\widetilde{D}$ into a versatile, well-oiled theory. This is the theory of *perfectoid spaces* [Sch12] [Sch13b].

[10] Pedantic note: in the case of a site, one has to replace intersections with fiber products.

Definition 5.7.5. Let C be a complete algebraically closed extension of $\mathbf{Q}_p$. A *perfectoid C-algebra* is a Banach C-algebra A such that the subring $A^\circ = \{a \in A | \sup \|a^n\| < \infty\}$ is bounded, and such that the Frobenius map

$$\begin{array}{ccc} A^\circ/pA^\circ & \to & A^\circ/pA^\circ \\ x & \mapsto & x^p \end{array}$$

is surjective.

Example 5.7.6. The simplest nontrivial example of a perfectoid C-algebra is $A = C\langle T^{1/p^\infty}\rangle$, equal to the completion of $C[T^{1/p^\infty}]$ with respect to the norm $\|\sum_r a_r T^r\| = \max |a_r|$. This is the ring of analytic functions on the perfectoid closed disc $\widetilde{D}$ of Example 5.7.4. It is not noetherian.

Given a perfectoid C-algebra A, one can define a *perfectoid affinoid space* $\operatorname{Spa} A$, a ringed space whose points are the continuous valuations[11] of A; Scholze shows that this is an *adic space* in the sense of Huber [Hub93]. Finally, a *perfectoid space* is an adic space created by gluing together perfectoid affinoid spaces.

A smooth rigid-analytic space X always admits a pro-étale cover $U_i \to X$, where each U_i is a perfectoid affinoid [Sch13a, Cor. 4.7].

Theorem 5.7.7 ([Sch12, Prop. 7.13])**.** *Let U be a perfectoid affinoid space over C. Then $H^i(U_{\text{ét}}, \mathcal{O}_U^+)$ is almost zero for $i > 0$.*

Here $\mathcal{O}_U^+$ is the sheaf of functions bounded by 1, considered as a sheaf on the étale site of U, and "almost zero" means annihilated by the maximal ideal of $\mathcal{O}_C = \{z \in C | \, |z| \leq 1\}$. This theorem implies that for a rigid-analytic space X, the cohomology $H^i_{\text{ét}}(X, \mathcal{O}_X^+)$ can "almost" be computed by the Čech complex associated to a pro-étale cover of X by perfectoid affinoids. Furthermore, the *Artin–Schreier sequence*

$$\begin{array}{ccccc} 0 \to \mathbf{Z}/p\mathbf{Z} \to \mathcal{O}_X^+/p & \to & \mathcal{O}_X^+/p \to 0 \\ x & \mapsto & x^p - x \end{array}$$

of sheaves on $X_{\text{ét}}$ can be used to deduce the same result for $H^i(X_{\text{ét}}, \mathbf{Z}/p\mathbf{Z})$.

For proper smooth X, the Artin–Schreier sequence is used to prove the following result (itself a stepping-stone to Theorem 5.6.2):

Theorem 5.7.8 ([Sch13a, Theorem 1.3])**.** *Let X be a proper smooth rigid-analytic space over C. Then $H^i(X_{\text{ét}}, \mathbf{Z}/p\mathbf{Z})$ is finite, and the map*

$$H^i(X_{\text{ét}}, \mathbf{Z}/p\mathbf{Z}) \otimes_{\mathbf{Z}/p\mathbf{Z}} \mathcal{O}_C/p \to H^i(X_{\text{ét}}, \mathcal{O}_X^+/p)$$

is an almost isomorphism of $\mathcal{O}_C/p$-modules, meaning that the kernel and cokernel are almost zero.

5.8. **Shimura varieties at infinite level.** The full force of the technology of perfectoid spaces is required for the proof of Theorem 5.5.1. In the context of that theorem, we are given an eigenclass g' in $H^i(\mathrm{Sh}, \mathbf{Z}/p\mathbf{Z})$. Using the comparison isomorphisms of (5.6.2), g' may be identified with a class in $H^i(\mathrm{Sh}^{\mathrm{an}}_{\text{ét}}, \mathbf{Z}/p\mathbf{Z})$, where $\mathrm{Sh}^{\mathrm{an}}$ is the rigid-analytic space over C associated with Sh.

[11] In contrast to the theory of rigid-analytic spaces, it is too naïve to consider only the maximal ideals of A.

Theorem 5.7.8 applies[12] to give an almost isomorphism

$$H^i(\mathrm{Sh}^{\mathrm{an}}_{\text{ét}}, \mathbf{Z}/p\mathbf{Z}) \otimes_{\mathbf{Z}/p\mathbf{Z}} \mathcal{O}_C/p \cong H^i(\mathrm{Sh}^{\mathrm{an}}_{\text{ét}}, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p).$$

Let g'' be the image of $g' \otimes 1$ under this isomorphism.

Because of Theorem 5.7.7, the cohomology $H^i_{\text{ét}}(\mathrm{Sh}^{\mathrm{an}}, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p)$ can be computed using the Čech complex associated to a pro-étale cover of $\mathrm{Sh}^{\mathrm{an}}$ consisting of perfectoid affinoid spaces. Astoundingly, there is a natural pro-étale cover coming from Shimura varieties *at infinite level*.

Theorem 5.8.1 ([Sch13c, Theorem I.8])**.** *For each $m \geq 0$, let Sh_{p^m} denote the Siegel modular variety parametrizing principally polarized abelian varieties of dimension n with Γ-level structure and an additional full p^m-level structure. Then the inverse limit $\varprojlim \mathrm{Sh}^{\mathrm{an}}_{p^m}$ exists as a perfectoid space $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$. There is a* period morphism

$$\pi\colon \mathrm{Sh}^{\mathrm{an}}_{p^\infty} \to \mathrm{Grass}(n, 2n)$$

onto the Grassmannian of n-planes in $2n$-space, having the following properties:

1. *π is affine, i.e., it pulls back affinoids to affinoids;*
2. *π commutes with the Hecke operators away from p;*
3. *letting $\mathcal{L}$ be the ample line bundle on $\mathrm{Grass}(n, 2n)$ coming from the Plücker embedding, $\pi^*\mathcal{L}$ is the line bundle ω from Example 5.4.2.*

Morally, the existence of π can be explained this way: A point of $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$ over C is an abelian variety A/C together with a basis for the Tate module T_pA. Taking duals, we get a trivialization $H^1(A_{\text{ét}}, \mathbf{Q}_p) \cong \mathbf{Q}_p^{2n}$. On the other hand we have the Hodge decomposition

$$H^1(A_{\text{ét}}, \mathbf{Q}_p) \otimes_{\mathbf{Q}_p} C \cong H^0(A, \Omega^1_{A/C}) \oplus H^1(A, \mathcal{O}_A)(-1).$$

Combining these structures, we get a distinguished n-plane $H^0(A, \Omega^1_{A/C})$ inside $H^1(A_{\text{ét}}, \mathbf{Q}_p) \otimes_{\mathbf{Q}_p} C \cong C^{2g}$, which is to say a point of $\mathrm{Grass}(n, 2n)$. It is quite another thing to show that $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$ is a perfectoid space; we can only refer the reader to [Sch13c] itself.

Let $s_1, \ldots, s_N$ be a basis for $H^0(\mathrm{Grass}(n, 2n), \mathcal{L})$. Then $\mathrm{Grass}(n, 2n)$ is covered by affinoid spaces $V_j = \{s_j \neq 0\}$. Each $U_j = \pi^{-1}(V_j)$ is an affinoid space by (1). The U_j constitute a covering of $\mathrm{Sh}^{\mathrm{an}}$ by perfectoid affinoid spaces in the pro-étale topology, and so $H^i(\mathrm{Sh}^{\mathrm{an}}_{\text{ét}}, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p)$ can be computed using the associated Čech complex. Crucially, each of the terms $H^0(U_J, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p)$ in that complex (here U_J is some intersection among the U_j) is stable under the prime-to-p Hecke operators, by (2). The existence of the eigenclass $g'' \in H^i(\mathrm{Sh}^{\mathrm{an}}_{\text{ét}}, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p)$ implies the existence of a function $g''' \in H^0(U_J, \mathcal{O}^+_{\mathrm{Sh}^{\mathrm{an}}}/p)$ with the same prime-to-p Hecke eigenvalues. This g''' can be thought of as the reduction mod p of a meromorphic function on $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$ with poles on a subspace of lower dimension.

Let $t_j = \pi^*(s_j)$, so t_j is a section of ω on $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$ (by property (3)). Then multiplying g''' by some sufficiently high power of $\prod_{j\in J} s_j$ produces a section $\overline{f}$ of $\omega^{\otimes k}$ mod p on $\mathrm{Sh}^{\mathrm{an}}_{p^\infty}$ for some k. This section is the reduction of a cusp form $f \in H^0(\mathrm{Sh}_{p^m}, \omega^{\otimes k})$, which is the cusp form required by Theorem 5.5.1.

[12]Actually one has to work with a compactified version of $\mathrm{Sh}^{\mathrm{an}}$ here, because Theorem 5.7.8 only applies to proper rigid-analytic spaces. We will be ignoring this issue for the purposes of exposition.

5.9. **Concluding remarks.** Theorem 5.2.1 is a spectacular advance. Even though special cases of it had been conjectured and tested numerically, nearly no one could have guessed that a proof was on the horizon. Nor was it clear, as it is now, that torsion classes in the cohomology of arithmetic manifolds play such an important role in number theory. (Even if one is only interested in a "characteristic 0" result like Theorem 5.1.1, Scholze's proof requires a detour through mod p^n cohomology.) We ought to mention Emerton's theory of completed cohomology (see [CE12] for a survey), which aims to establish a general theory of p-adic automorphic forms to complement the classical theory. As a byproduct of his proof of Theorem 5.2.1, Scholze shows that certain p-adic automorphic forms have corresponding Galois representations as well. These new Galois representations move in p-adic families and are not necessarily "geometric".

The diversity of methods required in the proof of Theorem 5.2.1 is remarkable. Prior theorems which constructed Galois representations from automorphic representations (e.g., [HT01]) required a combination of difficult techniques from automorphic forms (base change, endoscopic transfer) and from algebraic geometry (Shimura varieties, étale cohomology). To these techniques we must now add the advanced theory of p-adic analytic geometry (perfectoid spaces, the pro-étale topology).

Despite these advances, not even our original Question A has anything remotely like a complete solution. Such a solution could arrive in the form of Artin's conjecture or a generalized Serre's conjecture, concerning the modularity of Galois representations with coefficients in $\mathbf{C}$ or a finite field, respectively. As mentioned at the end of §4.5, Artin's conjecture is open except in dimension 1 and in some of the two-dimensional cases. The original *Serre's conjecture* [Ser87] refers to the modularity of an odd irreducible two-dimensional Galois representation with coefficients in a finite field; it is now a theorem [KW09a]. A generalized Serre's conjecture would be a converse to Theorem 5.2.1; this has been formulated precisely in [Her09] but for the moment remains wide open.

About the author

Jared Weinstein is an assistant professor of mathematics at Boston University with an interest in the Langlands program and arithmetic geometry. He received his Ph.D. from UC Berkeley.

Acknowledgments

This article is an expanded version of notes for the Current Events Bulletin session of the 2015 AMS-MAA Joint Mathematics Meetings. We thank David Eisenbud for the invitation to participate in that session. We also thank Keith Conrad for thoroughly proofreading a draft of the manuscript, and Sug Woo Shin and Ehud de Shalit for their comments. Finally, we thank the referee for an extraordinarily helpful and swift report.

References

[ADP02] Avner Ash, Darrin Doud, and David Pollack, *Galois representations with conjectural connections to arithmetic cohomology*, Duke Math. J. **112** (2002), no. 3, 521–579, DOI 10.1215/S0012-9074-02-11235-6. MR1896473 (2003g:11055)

[AG00] Avner Ash and Robert Gross, *Generalized non-abelian reciprocity laws: a context for Wiles' proof*, Bull. London Math. Soc. **32** (2000), no. 4, 385–397, DOI 10.1112/S0024609300007244. MR1760802 (2001h:11142)

[Ash92] Avner Ash, *Galois representations attached to mod p cohomology of* $\mathrm{GL}(n, \mathbf{Z})$, Duke Math. J. **65** (1992), no. 2, 235–255, DOI 10.1215/S0012-7094-92-06510-0. MR1150586 (93c:11036)

[BCDT01] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: wild 3-adic exercises*, J. Amer. Math. Soc. **14** (2001), no. 4, 843–939 (electronic), DOI 10.1090/S0894-0347-01-00370-8. MR1839918 (2002d:11058)

[BGR84] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis. A systematic approach to rigid analytic geometry*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 261, Springer-Verlag, Berlin, 1984. MR746961 (86b:32031)

[BJ06] Armand Borel and Lizhen Ji, *Compactifications of symmetric and locally symmetric spaces*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 2006. MR2189882 (2007d:22030)

[Blo81] Spencer Bloch, *Book Review: Étale cohomology*, Bull. Amer. Math. Soc. (N.S.) **4** (1981), no. 2, 235–239, DOI 10.1090/S0273-0979-1981-14894-1. MR1567311

[Bos11] Johan Bosman, *Polynomials for projective representations of level one forms*, Computational aspects of modular forms and Galois representations, Ann. of Math. Stud., vol. 176, Princeton Univ. Press, Princeton, NJ, 2011, pp. 159–172. MR2857091

[BS73] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Avec un appendice: Arrondissement des variétés à coins, par A, Douady et L. Hérault, Comment. Math. Helv. **48** (1973), 436–491. MR0387495 (52 #8337)

[Buh78] Joe P. Buhler, *Icosahedral Galois representations*, Lecture Notes in Mathematics, Vol. 654, Springer-Verlag, Berlin-New York, 1978. MR0506171 (58 #22019)

[BW00] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, 2nd ed., Mathematical Surveys and Monographs, vol. 67, American Mathematical Society, Providence, RI, 2000. MR1721403 (2000j:22015)

[Cas67] *Algebraic number theory*, Proceedings of an instructional conference organized by the London Mathematical Society (a NATO Advanced Study Institute) with the support of the International Mathematical Union. Edited by J. W. S. Cassels and A. Fröhlich, Academic Press, London; Thompson Book Co., Inc., Washington, D.C., 1967. MR0215665 (35 #6500)

[CE12] Frank Calegari and Matthew Emerton, *Completed cohomology—a survey*, Non-abelian fundamental groups and Iwasawa theory, London Math. Soc. Lecture Note Ser., vol. 393, Cambridge Univ. Press, Cambridge, 2012, pp. 239–257. MR2905536

[Clo90] Laurent Clozel, *Motifs et formes automorphes: applications du principe de fonctorialité* (French), Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 77–159. MR1044819 (91k:11042)

[Cox89] David A. Cox, *Primes of the form* $x^2 + ny^2$*. Fermat, class field theory and complex multiplication*, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1989. MR1028322 (90m:11016)

[Del71] Pierre Deligne, *Formes modulaires et représentations l-adiques* (French), Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363, Lecture Notes in Math., vol. 175, Springer, Berlin, 1971, pp. Exp. No. 355, 139–172. MR3077124

[Del79] Pierre Deligne, *Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques* (French), Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 247–289. MR546620 (81i:10032)

[DM06] Darrin Doud and Michael W. Moore, *Even icosahedral Galois representations of prime conductor*, J. Number Theory **118** (2006), no. 1, 62–70, DOI 10.1016/j.jnt.2005.08.008. MR2220262 (2006m:11077)

[DS74] Pierre Deligne and Jean-Pierre Serre, *Formes modulaires de poids* 1 (French), Ann. Sci. École Norm. Sup. (4) **7** (1974), 507–530 (1975). MR0379379 (52 #284)

[Eme06] Matthew Emerton, *A local-global compatibility conjecture in the p-adic Langlands programme for* $\mathrm{GL}_{2/\mathbb{Q}}$, Pure Appl. Math. Q. **2** (2006), no. 2, Special Issue: In honor of John H. Coates., 279–393, DOI 10.4310/PAMQ.2006.v2.n2.a1. MR2251474 (2008d:11133)

[Fal88] Gerd Faltings, *p-adic Hodge theory*, J. Amer. Math. Soc. **1** (1988), no. 1, 255–299, DOI 10.2307/1990970. MR924705 (89g:14008)

[Fal02] Gerd Faltings, *Almost étale extensions*, Astérisque **279** (2002), 185–270. Cohomologies p-adiques et applications arithmétiques, II. MR1922831 (2003m:14031)

[FM87] Jean-Marc Fontaine and William Messing, *p-adic periods and p-adic étale cohomology*, Current trends in arithmetical algebraic geometry (Arcata, Calif., 1985), Contemp. Math., vol. 67, Amer. Math. Soc., Providence, RI, 1987, pp. 179–207, DOI 10.1090/conm/067/902593. MR902593 (89g:14009)

[FM95] Jean-Marc Fontaine and Barry Mazur, *Geometric Galois representations*, Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), Ser. Number Theory, I, Int. Press, Cambridge, MA, 1995, pp. 41–78. MR1363495 (96h:11049)

[Fon94] Jean-Marc Fontaine, *Représentations p-adiques semi-stables* (French), Astérisque **223** (1994), 113–184. With an appendix by Pierre Colmez; Périodes p-adiques (Bures-sur-Yvette, 1988). MR1293972 (95g:14024)

[Fre86] Gerhard Frey, *Links between stable elliptic curves and certain Diophantine equations*, Ann. Univ. Sarav. Ser. Math. **1** (1986), no. 1, iv+40. MR853387 (87j:11050)

[Gel75] Stephen S. Gelbart, *Automorphic forms on adèle groups*, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1975. Annals of Mathematics Studies, No. 83. MR0379375 (52 #280)

[Gel84] Stephen Gelbart, *An elementary introduction to the Langlands program*, Bull. Amer. Math. Soc. (N.S.) **10** (1984), no. 2, 177–219, DOI 10.1090/S0273-0979-1984-15237-6. MR733692 (85e:11094)

[Gel97] Stephen Gelbart, *Three lectures on the modularity of* $\overline{\rho}_{E,3}$ *and the Langlands reciprocity conjecture*, Modular forms and Fermat's last theorem (Boston, MA, 1995), Springer, New York, 1997, pp. 155–207. MR1638479

[GJ72] Roger Godement and Hervé Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Mathematics, Vol. 260, Springer-Verlag, Berlin-New York, 1972. MR0342495 (49 #7241)

[Gor05] Mark Goresky, *Compactifications and cohomology of modular varieties*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 551–582. MR2192016 (2006h:14033)

[Hec27] E. Hecke, *Zur Theorie der elliptischen Modulfunktionen* (German), Math. Ann. **97** (1927), no. 1, 210–242, DOI 10.1007/BF01447866. MR1512360

[Her09] Florian Herzig, *The weight in a Serre-type conjecture for tame n-dimensional Galois representations*, Duke Math. J. **149** (2009), no. 1, 37–116, DOI 10.1215/00127094-2009-036. MR2541127 (2010f:11083)

[HLTT] M. Harris, K.-W. Lan, R. Taylor, and J. Thorne, *On the rigid cohomology of certain Shimura varieties*, Preprint.

[HT01] Michael Harris and Richard Taylor, *The geometry and cohomology of some simple Shimura varieties*, With an appendix by Vladimir G. Berkovich, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, NJ, 2001. MR1876802 (2002m:11050)

[Hub93] R. Huber, *Continuous valuations*, Math. Z. **212** (1993), no. 3, 455–477, DOI 10.1007/BF02571668. MR1207303 (94e:13041)

[Hub96] Roland Huber, *Étale cohomology of rigid analytic varieties and adic spaces*, Aspects of Mathematics, E30, Friedr. Vieweg & Sohn, Braunschweig, 1996. MR1734903 (2001c:14046)

[IR90] Kenneth Ireland and Michael Rosen, *A classical introduction to modern number theory*, 2nd ed., Graduate Texts in Mathematics, vol. 84, Springer-Verlag, New York, 1990. MR1070716 (92e:11001)

[Kis09] Mark Kisin, *The Fontaine-Mazur conjecture for* GL_2, J. Amer. Math. Soc. **22** (2009), no. 3, 641–690, DOI 10.1090/S0894-0347-09-00628-6. MR2505297 (2010j:11084)

[KM85] Nicholas M. Katz and Barry Mazur, *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies, vol. 108, Princeton University Press, Princeton, NJ, 1985. MR772569 (86i:11024)

[Kob84] Neal Koblitz, *Introduction to elliptic curves and modular forms*, Graduate Texts in Mathematics, vol. 97, Springer-Verlag, New York, 1984. MR766911 (86c:11040)

[KW09a] Chandrashekhar Khare and Jean-Pierre Wintenberger, *On Serre's conjecture for 2-dimensional mod p representations of* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Ann. of Math. (2) **169** (2009), no. 1, 229–253, DOI 10.4007/annals.2009.169.229. MR2480604 (2009m:11077)

[KW09b] Chandrashekhar Khare and Jean-Pierre Wintenberger, *Serre's modularity conjecture. I*, Invent. Math. **178** (2009), no. 3, 485–504, DOI 10.1007/s00222-009-0205-7. MR2551763 (2010k:11087)

[Laf02] Laurent Lafforgue, *Chtoucas de Drinfeld et correspondance de Langlands* (French, with English and French summaries), Invent. Math. **147** (2002), no. 1, 1–241, DOI 10.1007/s002220100174. MR1875184 (2002m:11039)

[Lan80] Robert P. Langlands, *Base change for* GL(2), Annals of Mathematics Studies, vol. 96, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980. MR574808 (82a:10032)

[Lan94] Serge Lang, *Algebraic number theory*, 2nd ed., Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994. MR1282723 (95f:11085)

[Mar91] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, Berlin, 1991. MR1090825 (92h:22021)

[Mil80] James S. Milne, *Étale cohomology*, Princeton Mathematical Series, vol. 33, Princeton University Press, Princeton, N.J., 1980. MR559531 (81j:14002)

[Mil13] James S. Milne, *Lie algebras, algebraic groups, and Lie groups*, 2013, Available at www.jmilne.org/math/.

[Neu99] Jürgen Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher; With a foreword by G. Harder. MR1697859 (2000m:11104)

[PS73] I. I. Pjateckii-Sapiro, *Zeta-functions of modular curves*, Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Springer, Berlin, 1973, pp. 317–360. Lecture Notes in Math., Vol. 349. MR0337975 (49 #2744)

[Rag04] M. S. Raghunathan, *The congruence subgroup problem*, Proc. Indian Acad. Sci. Math. Sci. **114** (2004), no. 4, 299–308, DOI 10.1007/BF02829437. MR2067695 (2005g:20081)

[Ram00] Ravi Ramakrishna, *Infinitely ramified Galois representations*, Ann. of Math. (2) **151** (2000), no. 2, 793–815, DOI 10.2307/121048. MR1765710 (2001e:11057)

[Rib90] 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, DOI 10.1007/BF01231195. MR1047143 (91g:11066)

[Sch12] Peter Scholze, *Perfectoid spaces*, Publ. Math. Inst. Hautes Études Sci. **116** (2012), 245–313, DOI 10.1007/s10240-012-0042-x. MR3090258

[Sch13a] Peter Scholze, *p-adic Hodge theory for rigid-analytic varieties*, Forum Math. Pi **1** (2013), e1, 77, DOI 10.1017/fmp.2013.1. MR3090230

[Sch13b] Peter Scholze, *Perfectoid spaces: a survey*, Current developments in mathematics 2012, Int. Press, Somerville, MA, 2013, pp. 193–227. MR3204346

[Sch13c] Peter Scholze, *Torsion in the cohomology of locally symmetric spaces*, Preprint, Bonn, 2013.

[Sen14] M. H. Sengün, *Arithmetic aspects of Bianchi groups*, Computations with Modular Forms: Proceedings of a summer school and conference, Heidelberg, August/September 2011, Contributions in Mathematical and Compuational Sciences, vol. 6, Springer, 2014, pp. 279–315.

[Ser73] J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973. Translated from the French; Graduate Texts in Mathematics, No. 7. MR0344216 (49 #8956)

[Ser87] Jean-Pierre Serre, *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ (French), Duke Math. J. **54** (1987), no. 1, 179–230, DOI 10.1215/S0012-7094-87-05413-5. MR885783 (88g:11022)

[Shi66] Goro Shimura, *A reciprocity law in non-solvable extensions*, J. Reine Angew. Math. **221** (1966), 209–220. MR0188198 (32 #5637)

[SRY12] R. Sujatha, H. N. Ramaswamy and C. S. Yogananda (editors), *Math unlimited*, Essays in mathematics, Science Publishers, Enfield, NH; distributed by CRC Press, Boca Raton, FL, 2012. MR2885277 (2012i:00002)

[Tat67] J. T. Tate, *p-divisible groups*, Proc. Conf. Local Fields (Driebergen, 1966), Springer, Berlin, 1967, pp. 158–183. MR0231827 (38 #155)

[Tat71] John Tate, *Rigid analytic spaces*, Invent. Math. **12** (1971), 257–289. MR0306196 (46 #5323)

[Tun81] Jerrold Tunnell, *Artin's conjecture for representations of octahedral type*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), no. 2, 173–175, DOI 10.1090/S0273-0979-1981-14936-3. MR621884 (82j:12015)

[TW95] Richard Taylor and Andrew Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2) **141** (1995), no. 3, 553–572, DOI 10.2307/2118560. MR1333036 (96d:11072)

[Was97] Lawrence C. Washington, *Introduction to cyclotomic fields*, 2nd ed., Graduate Texts in Mathematics, vol. 83, Springer-Verlag, New York, 1997. MR1421575 (97h:11130)

[Wei82] André Weil, *Adeles and algebraic groups*, Progress in Mathematics, vol. 23, Birkhäuser, Boston, Mass., 1982. With appendices by M. Demazure and Takashi Ono. MR670072 (83m:10032)

[Wil95] Andrew Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551, DOI 10.2307/2118559. MR1333035 (96d:11071)

[Wym72] B. F. Wyman, *What is a reciprocity law?*, Amer. Math. Monthly **79** (1972), 571–586; correction, ibid. 80 (1973), 281. MR0308084 (46 #7199)

Department of Mathematics, Boston University, Boston, Massachusetts