Exploring the Galois group of the rational numbers: recent breakthroughs

Jared Weinstein

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.

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, given a prime p, determines whether $f(x)$ is split modulo p?

This motivating question is lifted almost verbatim from B. Wyman's 1972 article, [Wym72], of which the present article is merely an updated version. It may surprise the reader to learn that 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 will be 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 three successive epochs:

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. 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 *abelian*. They only apply to those instances of Question A where the polynomial $f(x)$ is solvable.

2. In the second half of the 20th century, a remarkable link was found between *modular forms* and 2-dimensional representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, due to Eichler, Shimura, and especially Deligne. This made it possible to find reciprocity laws for certain quintic $f(x)$ which are not solvable.

3. 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 [Sch13], which employs techniques invented within the past three years.

2 Fermat, Gauss, and solvable 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.

Thus we are reduced to the case that $n = p$ is prime. We wish 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 $x^2 + 1 \equiv 0 \pmod{p}$. Solving it is equivalent to Question A for $f(x) = x^2 + 1$.

Theorem 2.1. *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 by Wilson's theorem is $\equiv -1 \pmod{p}$. □

Another way of phrasing Thm. 2.1 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 $f(x)$, 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.)

Thm. 2.1 demonstrates the simplest possible sort of reciprocity law, which namely one where the factorization of $f(x)$ modulo p is determined by a *congruence condition on p.* This is also the case for x^2+x+1, which splits modulo p if and only if $p \equiv 1 \pmod 3$. (Sketch of proof: if $p \equiv 1 \pmod 3$, and g is a generator of the cyclic group $(\mathbf{Z}/p\mathbf{Z})^\times$, then $x = g^{(p-1)/3}$ is a root.) In fact one has a congruence condition whenever $f(x)$ is a quadratic polynomial:

Theorem 2.2 (Quadratic Reciprocity)**.** *Let $f(x) = x^2 + bx + c$ be an irreducible polynomial, so that $d = b^2 - 4c$ 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.*

Note that $f(x)$ factors modulo p if and only if d is congruent to a square modulo p. One introduces the *Legendre symbol* $\left(\frac{d}{p}\right)$ for any odd prime p and any integer d prime to p, defined to be 1 if d is a square modulo p and -1 otherwise. Thus for instance Thm. 2.1 is the statement that $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$. In elementary number theory texts one learns a more precise version of Thm. 2.2: 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? Thm. 2.2 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 primes (for instance 7) are 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 be a principal ideal domain. 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].

What about polynomials $f(x)$ of higher degree? A little experimentation will reveal that the factorization behavior of a "random" cubic will be influenced, but not completely determined by, a congruence condition modulo p. There are special cases: for instance the polynomial $x^3 + x^2 - 2x + 1$ splits modulo p if and only if $p \equiv \pm 1 \pmod 7$. So when is the splitting behavior of a polynomial determined by congruence conditions?

For a clue, let $m \geq 1$ and consider the polynomial $x^m - 1$. It 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 *cyclotomic polynomial*, that is, an irreducible divisor of some $x^m - 1$. Using some algebraic number theory, one even gets congruence conditions for those $f(x)$ whose roots are contained in a *cyclotomic field* $\mathbf{Q}(\zeta_m)$, where $\zeta_m = \exp(2\pi i/m)$. (The roots of $x^3 + x^2 - 2x + 1$, for instance, are $2\cos(2\pi k/7)$, where $k = 1, 2, 3$.)

What would a reciprocity law look like if it isn't a congruence condition? As with the quadratic reciprocity law, the following theorem was conjectured by Euler and proved by Gauss.

Theorem 2.3. *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 $a^2 + b^2$, the representation of p by the quadratic form $a^2 + 64b^2$ is not determined by a congruence condition on p. But in fact there is a disguised congruence condition in Thm. 2.3, which was well known to Gauss. 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 Thm. 2.1 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 $p = \omega\overline{\omega}$, where $\omega = a + bi$. Thm. 2.3 says that $x^4 - 2$ splits modulo p if and only if ω is congruent to a rational integer modulo 8. Indeed, this condition translates into the statement that $b = 8b_0$ for an integer b_0, in which case $p = a^2 + 64b_0^2$. 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.

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 into prime elements, it does have the corresponding property for *ideals*. As an example, in the ring $\mathbf{Z}[\sqrt{-5}]$, the element 6 admits the two factorizations $2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ into irreducible elements, none of which divide any other. However, the ideal $6\mathcal{O}_K$ admits a factorization into prime ideals in one way only: $6\mathcal{O}_K = (2, 1 + \sqrt{-5})^2(3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5})$.

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 prime ideals. In fact we have $e_i = 1$ for all i unless p belongs to a finite list of *ramified primes*. For each i, the quotient $\mathcal{O}_K/\mathfrak{p}_i$ is a finite field extension of $\mathbf{F}_p$; the *degree* of $\mathfrak{p}_i$ is defined as the degree of this extension. We say that p is *split* in K if $p\mathcal{O}_K$ is the product of distinct prime ideals of degree 1. These concepts all have relative notions with respect to an extension of number fields L/K.

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 normal (meaning that it is the splitting field of a collection of polynomials with coefficients in K) and separable (meaning that the minimal polynomial of any element of L over K has no repeated roots). If L/K is Galois, one defines the Galois group $\mathrm{Gal}(L/K)$ as 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.

As an example, the splitting field of the polynomial $x^4 - 2$ over $\mathbf{Q}$ is $L = \mathbf{Q}(i, 2^{1/4})$. The Galois group $\mathrm{Gal}(L/\mathbf{Q})$ is the dihedral group of order 8, generated by two elements σ and τ, defined by the table

$$\begin{aligned} \sigma(2^{1/4}) &= i2^{1/4}, & \tau(2^{1/4}) &= 2^{1/4}, \\ \sigma(i) &= i, & \tau(i) &= i. \end{aligned}$$

These generators satisfy the relations $\sigma^4 = 1$, $\tau^2 = 1$, and $\tau\sigma\tau = \sigma^{-1}$.

If L/K is Galois and $\mathfrak{p}$ is a prime ideal of K which is unramified in L, let $\mathfrak{P}$ be a prime ideal of L dividing $\mathfrak{p}$. The number of elements of $\mathcal{O}_K/\mathfrak{p}$ is denoted $N\mathfrak{p}$. The Galois group $\mathrm{Gal}((\mathcal{O}_L/\mathfrak{P})/(\mathcal{O}_K/\mathfrak{p}))$ is cyclic of order equal to the degree of $\mathfrak{P}|\mathfrak{p}$, with a distinguished generator $x \mapsto x^{N\mathfrak{p}}$. It turns out that there exists a unique element $\mathrm{Frob}_{\mathfrak{P}|\mathfrak{p}} \in \mathrm{Gal}(L/K)$, the *Frobenius*

element, which lifts this generator. That is:

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

for all $x \in \mathcal{O}_L$. If a different prime $\mathfrak{P}'$ dividing $\mathfrak{p}$ is chosen, the resulting Frobenii $\mathrm{Frob}_{\mathfrak{P}}$ and $\mathrm{Frob}_{\mathfrak{P}'}$ are conjugate in $\mathrm{Gal}(L/K)$. Thus one can talk about $\mathrm{Frob}_{\mathfrak{p}}$ as a well-defined *conjugacy class* in $\mathrm{Gal}(L/K)$. An important observation is that

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

Class field theory refers to the complete solution of Question B in the case that $\mathrm{Gal}(L/K)$ is *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 $\mathfrak{p}$, where the modulus is an ideal $\mathfrak{f}$ of K divisible only by ramified primes. Rather than making this precise, we spell out the example relevant to Gauss' $a^2 + 64b^2$ theorem.

Example 2.4. Let $K = \mathbf{Q}(i)$ and $L = K(2^{1/4})$, so that $\mathrm{Gal}(L/K)$ is a cyclic group of order 4 generated by σ. Here $\mathfrak{f} = (8)$. Class field theory shows that if $\mathfrak{p} = (\omega)$ is a prime of K which is relatively prime to 2, then $\mathrm{Frob}_{\mathfrak{p}}$ is determined by the image of ω in $(\mathbf{Z}[i]/8\mathbf{Z}[i])^\times$, in the following way. The group $(\mathbf{Z}[i]/8\mathbf{Z}[i])^*$ is generated by the subgroup $(\mathbf{Z}/8\mathbf{Z})^\times$ and the elements i and $1 + 2i$. There exists a unique surjective homomorphism

$$r\colon (\mathbf{Z}[i]/8\mathbf{Z}[i])^\times \to \mathrm{Gal}(L/K)$$

which is trivial on $(\mathbf{Z}/8\mathbf{Z})^\times$ and i, and which sends $1 + 2i$ to σ. Then $\mathrm{Frob}_{\mathfrak{p}} = r(\omega)$. As a result, $\mathrm{Frob}_{\mathfrak{p}} = 1$ if and only if $\mathfrak{p} = (p)$ for $p \equiv 3 \pmod 4$ or else if $\mathfrak{p} = (a + 8bi)$ with $p = a^2 + 64b^2$ prime.

Class field theory 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 p splits in K_1, a prime above p in K_1 splits in K_2, and so on, with each splitting being governed by congruences. In Example 2.4, the relevant tower was $\mathbf{Q} \subset \mathbf{Q}(i) \subset \mathbf{Q}(i, 2^{1/4})$.

It is immensely useful to talk about all of the extensions of $\mathbf{Q}$ at once, as living in an algebraic closure $\overline{\mathbf{Q}}$. One considers the *absolute Galois group* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$; this is just the automorphism group of the field $\overline{\mathbf{Q}}$. More to the point, we have

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) = \varprojlim_K \mathrm{Gal}(K/\mathbf{Q}),$$

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

Example 2.5. The group D_8 has a two-dimensional representation which sends σ to $\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ and τ to $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$. The character of this representation is 2 on the identity of D_8, -2 on σ^2, and 0 everywhere else. Thus we can construct a 2-dimensional Galois representation

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

which factors through $\mathrm{Gal}(\mathbf{Q}(i, 2^{1/4})/\mathbf{Q})$. This will have the property that for all odd primes p, ρ is unramified at p, meaning that the fixed field of the kernel of ρ is unramified at p. Consequently $\rho(\mathrm{Frob}_p)$ is well defined. We have

$$\mathrm{tr}\,\rho(\mathrm{Frob}_p) = \begin{cases} 2, & p = a^2 + 64b^2, \\ -2, & p = a^2 + 16b^2,\ b \text{ odd}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

3 Elliptic modular forms

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 Abel and Jacobi in the early 19th century, which in turn arose in association with finding the arc length of an ellipse. For an introduction to the subject, we recommend the book [Ser73].

In brief, 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 a hyperbolic metric. The automorphism group of $\mathcal{H}$ is the group of *Möbius transformations* $z \mapsto (az+b)/(cz+d)$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{R})$. In brief, a modular form is a holomorphic function $f(\tau)$ on $\mathcal{H}$ which transforms in a certain way under a subgroup of $\mathrm{SL}_2(\mathbf{R})$.

For a nonzero integer N, let $\Gamma_0(N)$ denote the subgroup of $\mathrm{SL}_2(\mathbf{Z})$ consisting of matrices which are upper-triangular modulo N.

Definition 3.1. Let $N, k \geq 1$ be integers, and let $\chi\colon (\mathbf{Z}/N\mathbf{Z})^\times \to \mathbf{C}^\times$ be a homomorphism. A *modular form of weight k, level N and character χ* is a holomorphic function g on $\mathcal{H}$ which satisfies

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

for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, and which is "holomorphic at the cusps".

In particular $g(\tau+1) = g(\tau)$, 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 all functions $g((a\tau+b)/(c\tau+d))$ with $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$. We say g is a *cusp form* if it is zero at the cusps, meaning that $a_0(g) = 0$ as well.

Some modular forms known as *theta functions* arise from sums involving rings of integers in quadratic fields, such as $\mathbf{Z}[i]$. Suppose $\mathfrak{f} \subset \mathbf{Z}[i]$ is an ideal, and $\chi\colon (\mathbf{Z}[i]/\mathfrak{f})^\times \to \mathbf{C}^\times$ is a nontrivial homomorphism. Extend χ to a function on $\mathbf{Z}[i]$ by declaring it 0 on elements which are not prime to $\mathfrak{f}$. Then

$$\theta_\chi(\tau) = \frac{1}{4}\sum_{\alpha\in\mathbf{Z}[i]} \chi(\alpha)q^{N(\alpha)}$$

is a modular form of weight 1 and level $4N(\mathfrak{f})$.

Example 3.2. Let $\chi\colon (\mathbf{Z}[i]/8\mathbf{Z}[i])^\times \to \mathbf{C}^\times$ be the homomorphism which is trivial on i and $(\mathbf{Z}/8\mathbf{Z})^\times$, and which sends $1+2i$ to i. Then θ_χ is a modular form of weight 1; for a prime p, its pth Fourier coefficient is

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

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 Eq. (2.1), we find that

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

for the Galois representation ρ constructed in Example 2.5. This equation hints at an extraordinary relationship between modular forms and representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

The space of cusp forms of weight k and level N is finite-dimensional. For each prime p not dividing N one defines a *Hecke operator* T_p on this space, which has the following effect on q-expansions:

$$T_p g(\tau) = \sum_{n \geq 0} a_{pn}(g) q^n + p^{k-1} \sum_{n \geq 0} a_n(g) q^{pn}.$$

(There are similar operators for primes dividing N.) These operators commute with one another, and so it makes sense to attempt to diagonalize them simultaneously. A modular form is an *eigenform* if it is an eigenvector for all Hecke operators. If $g = \sum_{n \geq 1} a_n(g) q^n$ is a cuspidal eigenform with $a_1(g) = 1$, then the eigenvalue of T_p on g is just $a_p(g)$.

Theorem 3.3. *Let $g(\tau) = \sum_{n \geq 1} a_n(g) q^n$ be a cuspidal eigenform of weight k and level N with character χ. Let E a number field containing the $a_n(g)$.*

1. *Suppose $k \geq 2$. Then for all prime ideals λ of $\mathcal{O}_E$ there exists an odd irreducible Galois representation*

$$\rho_{g,\lambda} \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathcal{O}_{E,\lambda})$$

   such that for all p prime to $N\lambda$, $\rho_{g,\lambda}$ is unramified at p, and the characteristic polynomial of $\rho_{g,\lambda}(\mathrm{Frob}_p)$ is $x^2 - a_p(g)x + \chi(p)p^{k-1}$.

2. *Suppose $k = 1$. Then there exists an odd irreducible Galois representation*

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

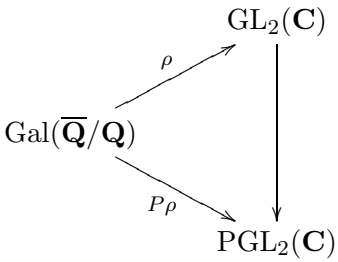
   such that for all p prime to N, ρ_g is unramified at p, and the characteristic polynomial of $\rho_g(\mathrm{Frob}_p)$ is $x^2 - a_p(g)x + \chi(p)$.

These two statements are proved in [Del71] and [DS74], respectively. In the first statement, $\mathcal{O}_{E,\lambda}$ is the completion of $\mathcal{O}_E$ with respect to the ideal λ; the image of $\rho_{g,\lambda}$ is infinite. In the second statement, where $k = 1$, the image of ρ_g is finite. A Galois representation ρ is *odd* if $\det \rho(c) = -1$, where c is complex conjugation.

Example 3.4 (An icosahedral form)**.** The following example is due to Joe Bühler, [Buh78]. Let

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

The discriminant of $f(x)$ is $2^6 5^8 11^2$, a square number. This means that the Galois group of f is contained in the icosahedral group A_5; in fact it equals A_5. The group A_5 doesn't have any irreducible 2-dimensional representations, but there exists a 4-fold cover $\widetilde{A}_5$ which does. It can be shown that there is an odd irreducible representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ whose image is $\widetilde{A}_5$, such that in the diagram



the fixed field of the kernel of $P\rho$ is the splitting field K of f. Artin's conjecture predicts a weight 1 cusp form $g(\tau)$ of level 800 and character χ associated to ρ, where χ is a character of conductor 100 and order 10. Indeed there is one:

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

where $i = \sqrt{-1}$ and $j = (1+\sqrt{5})/2$. A prime $p \neq 2, 5$ splits in K if and only if $\rho(\mathrm{Frob}_p)$ is a scalar matrix. Since $\rho(\mathrm{Frob}_p)$ has finite order, it is semisimple, and therefore it is scalar if and only if its characteristic polynomial has zero discriminant. But the characteristic polynomial is $x^2 - a_p(g)x + \chi(p)$, with discriminant $a_p(g)^2 - 4\chi(p)$. Therefore we have the following answer to Question B: p splits in K if and only if $a_p(g)^2 = 4\chi(p)$.

Example 3.5 (The Ramanujan Δ-function)**.** The product

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

defines a cuspidal eigenform of weight 12 and level 1, and so Thm. 3.3 associates to it an ℓ-adic representation $\rho_{\Delta,\ell}$ for all primes ℓ. This can be reduced modulo ℓ to obtain a mod ℓ Galois representation $\overline{\rho}_{\Delta,\ell}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to$

$\mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$, whose kernel cuts out a number field which is ramified only at ℓ. It is a difficult computational problem to compute this number field. For some small primes ℓ this has been carried out in [Bos11], at least for the associated projective representation $P\overline{\rho}_{\Delta,\ell}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{PGL}_2(\mathbf{Z}/\ell\mathbf{Z})$. For instance if $\ell = 11$, the fixed field of the kernel of $P\overline{\rho}_{\Delta,\ell}$ is the splitting field of

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

From this we can conclude the following reciprocity law, valid for almost all p: if $f(x)$ splits modulo p then $\tau(p)^2 \equiv 4p \pmod{11}$.

4 The cohomology of arithmetic manifolds

Modular forms are holomorphic forms on $\mathcal{H}$ which admit symmetries with respect to a finite-index subgroup $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbf{Z})$. It stands to reason that they correspond to objects defined on the quotient $Y_0(N) = \Gamma_0(N)\backslash\mathcal{H}$, a (non-compact) Riemann surface. For instance, if $f(\tau)$ is a modular form of weight 2, level N, and trivial character, then $f(\tau)d\tau$ is *invariant* under $\Gamma_0(N)$, and so descends to a differential form on $Y_0(N)$. If $f(\tau)$ happens to be a cusp form, then $f(\tau)d\tau$ extends to a differential form on $X_0(N)$, the smooth compactification of $Y_0(N)$. In fact the space of cusp forms of weight 2 is isomorphic to the space $H^0(X_0(N), \Omega^1_{X_0(N)/\mathbf{C}})$ of holomorphic differential forms on $X_0(N)$.

On the other hand, the Hodge decomposition for the compact Riemann surface $X_0(N)$ shows that the singular cohomology $H^1(X_0(N), \mathbf{C})$ is the direct sum of $H^0(X_0(N), \Omega^1_{X_0(N)/\mathbf{C}})$ and its complex conjugate. All of these spaces come equipped with actions by the Hecke operators T_p. The conclusion is that systems of Hecke eigenvalues coming from weight 2 forms are present already in the singular cohomology $H^1(X_0(N), \mathbf{C})$. (There is a similar statement for forms of higher weight; one replaces the $\mathbf{C}$ in $H^1(X_0(N), \mathbf{C})$ with a non-constant coefficient system.) Therefore one could have phrased Thm. 3.3 (at least the part pertaining to forms of weight $k \geq 2$) in terms of *Hecke eigenclasses* in the singular cohomology of $X_0(N)$. (Equivalently, one can phrase it in terms of the group cohomology $H^1(\Gamma_0(N), \mathbf{C})$.)

One might seek to generalize Thm. 3.3 to higher dimension as follows. The upper halfplane $\mathcal{H}$ is the quotient $\mathrm{SL}_2(\mathbf{R})/\mathrm{SO}(2)$, so let us put $\mathcal{H}_n = \mathrm{SL}_n(\mathbf{R})/\mathrm{SO}(n)$; this is a manifold with a left action by $\mathrm{SL}_n(\mathbf{R})$. Let

us abuse notation and write $\Gamma_0(N) \subset \mathrm{SL}_n(\mathbf{Z})$ for the subgroup of matrices whose first column is congruent to $(*, 0, \ldots, 0)$ modulo N. Then one can form the quotient $\Gamma_0(N)\backslash\mathcal{H}_n$, an *arithmetic manifold*. The cohomology $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{C})$ admits actions by Hecke operators. For each prime $p \nmid N$, there isn't just one Hecke operator T_p but rather $n-1$ operators $T_{p,1}, \ldots, T_{p,n-1}$. These operators commute with one another, and so one can talk about *eigenclasses* in $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{C})$ for all the $T_{p,i}$. Do these correspond to n-dimensional Galois representations?

The main obstacle to generalizing Thm. 3.3 is this: in the $n = 2$ case, $\mathcal{H} = \mathcal{H}_2$ is a complex manifold and $X_0(N)$ is an algebraic curve which even admits a model over the rational numbers. This fact is critical for the construction of Galois representations, which live in the ℓ-adic étale cohomology of $X_0(N)$. However if $n > 2$, $\mathcal{H}_n$ isn't even a complex manifold, and so no quotient of it is going to be an algebraic variety. (For instance, $\mathcal{H}_3$ has dimension 5, which is odd.) Nonetheless, the following theorem was announced in 2012:

Theorem 4.1 ([HLTT],[Sch13])**.** *Let g be a Hecke eigenclass in the singular cohomology $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{C})$, and let $a_{p,i}(g)$ be the eigenvalue of $T_{p,i}$ on g for $p \nmid N$ prime. Let E be a number field containing all the $a_{p,i}(g)$, and let λ be a prime of E. Then there exists a continuous semisimple Galois representation*

$$\rho_{g,\lambda} \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(E_\lambda)$$

which is associated to g in the sense that for all primes p which are prime to $N\lambda$, $\rho_{g,\lambda}$ is unramified at Frob_p*, and the characteristic polynomial of* $\rho_{g,\lambda}(\mathrm{Frob}_p)$ *is*

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

The results of [HLTT] and [Sch13] 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". It would take us rather far afield to define the terms in the preceding sentence, but suffice it to say that Thm. 4.1 is the special case $F = \mathbf{Q}$.

In fact the results of [Sch13] are stronger still. Thm. 4.1 concerns the singular cohomology $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{C})$ with complex coefficients, but we could also have considered the integral cohomology $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{Z})$, a finitely generated abelian group equipped with the action of Hecke operators $T_{p,i}$.

When $n = 2$, $Y_0(N) = \Gamma_0(N)\backslash\mathcal{H}$ is a surface; the integral cohomology groups of a surface are known to be torsion-free. But for $n > 2$, the cohomology $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \mathbf{Z})$ can contain a large torsion subgroup. This torsion subgroup is also preserved by the Hecke operators, and one may ask whether the *torsion* eigenclasses have corresponding *torsion* Galois representations. In fact they do:

Theorem 4.2 ([Sch13])**.** *Let ℓ be prime, and let g be a Hecke eigenclass in $H^j(\Gamma_0(N)\backslash\mathcal{H}_n, \overline{\mathbf{F}}_\ell)$. Then there exists a continuous semisimple Galois representation*

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

which is associated to g in the sense of Thm. 4.1.

In prior years, Ash and others had developed Serre-type conjectures which predict that every Galois representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_3(\overline{\mathbf{F}}_\ell)$ has a corresponding Hecke eigenclass g in $H^j(\Gamma_0(N), V)$ for an appropriate choice of N, j, and an $\overline{\mathbf{F}}_\ell[\mathrm{SL}_3(\mathbf{Z})]$-module V. See for instance [ADP02], which offers a great deal of numerical evidence in the form of pairs (ρ, g), where ρ and g appear to be associated in the sense that the characteristic polynomial of $\rho(\mathrm{Frob}_p)$ is as in Thm. 4.1 for the first few primes p. Thm. 4.2 shows that there really is a ρ' attached to each g, and then examination of sufficiently many small primes is enough to prove that $\rho = \rho'$. In those cases one has a reciprocity law for the fixed field K of $\ker\rho$: the splitting behavior of primes in K is governed by the Hecke eigenvalues of the eigenclass g.

Thm. 4.2 is truly spectacular. It links Galois representations with torsion classes in the cohomology of arithmetic manifolds, which don't necessarily come from automorphic representations (these had been the starting point for most generalizations of Thm. 3.3). The method of proof is striking. The first step, an idea suggested by Clozel, is to show that the arithmetic manifold $\Gamma_0(N)\backslash\mathcal{H}_n$ appears "at the boundary" of a Shimura variety Sh_N, which implies that an eigenclass g as in Thm. 4.2 shows up as an eigenclass in the cohomology $H^i(\mathrm{Sh}_N, \overline{\mathbf{F}}_\ell)$. The next step is to show that there exists a cuspidal eigenform on some higher level Shimura variety $\mathrm{Sh}_{N\ell^m}$ whose mod ℓ eigenvalues match those of g. This required working with a Shimura variety $\mathrm{Sh}_{N\ell^\infty}$ *at infinite level.*

The space $\mathrm{Sh}_{N\ell^\infty}$ isn't an algebraic variety or even a manifold. Rather, it is a fractal-like entity known as a *perfectoid space.* Perfectoid spaces were also devised by Scholze in [Sch12] for completely different ends; in this application, Scholze proves a comparison theorem for perfectoid spaces which

links mod ℓ étale cohomology and coherent cohomology. This comparison theorem furnishes the required cusp form, and with it the Galois representation.

Despite these remarkable advances, there are still major unsolved problems in our search for reciprocity laws. We conclude with a list of open questions.

- By Thm. 3.3, modular forms of weight 1 correspond to *odd* 2-dimensional Galois representations $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$. The general philosophy of Langlands predicts that *even* 2-dimensional Galois representations ought to correspond to algebraic Maass forms. A *Maass form* is an analytic (not holomorphic) function on $\Gamma_0(N)\backslash\mathcal{H}$ which is an eigenvector for the Laplacian $-y^{-2}(\partial^2/\partial x^2 + \partial^2/\partial y^2)$; it is *algebraic* if the eigenvalue is $1/4$. Nobody has any idea how the correspondence works in either direction, outside of the "solvable" cases.

- If $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ is an irreducible Galois representation, subject to a suitable condition at ℓ, must it arise from an eigenclass g as in Thm. 4.1? This is a generalization of the Fontaine-Mazur conjecture, [FM95], which for $n = 2$ was proved by Kisin, [Kis09], save some exceptional cases. This is a question of *modularity* of Galois representations, of which there is a large amount of literature. We remark in passing that a special case of modularity was key to Wiles' attack on Fermat's Last Theorem.

- If $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{F}}_\ell)$ is an irreducible Galois representation, must it arise from an eigenclass g as in Thm. 4.2? The case of $n = 2$ and ρ odd is *Serre's conjecture*, proved by Khare and Wintenberger, [KW09].

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. MR 1896473 (2003g:11055)

[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. MR 2857091

[Buh78] Joe P. Buhler, *Icosahedral Galois representations*, Lecture Notes in Mathematics, Vol. 654, Springer-Verlag, Berlin-New York, 1978. MR 0506171 (58 #22019)

[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. MR 1028322 (90m:11016)

[Del71] Pierre Deligne, *Formes modulaires et représentations l-adiques*, Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363, Lecture Notes in Math., vol. 175, Springer, Berlin, 1971, pp. Exp. No. 355, 139–172. MR 3077124

[DS74] Pierre Deligne and Jean-Pierre Serre, *Formes modulaires de poids 1*, Ann. Sci. École Norm. Sup. (4) **7** (1974), 507–530 (1975). MR 0379379 (52 #284)

[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. MR 1363495 (96h:11049)

[HLTT] Michael Harris, Kai-Wen Lan, Richard Taylor, and Jack Thorne, *On the rigid cohomology of certain Shimura varieties*, Preprint.

[Kis09] Mark Kisin, *The Fontaine-Mazur conjecture for* GL_2, J. Amer. Math. Soc. **22** (2009), no. 3, 641–690. MR 2505297 (2010j:11084)

[KW09] 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. MR 2480604 (2009m:11077)

[Sch12] Peter Scholze, *Perfectoid spaces*, Publ. Math. Inst. Hautes Études Sci. **116** (2012), 245–313. MR 3090258

[Sch13] ———, *Torsion in the cohomology of locally symmetric spaces*, Preprint, Bonn, 2013.

[Ser73] J.-P. Serre, *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7. MR 0344216 (49 #8956)

[Wym72] B. F. Wyman, *What is a reciprocity law?*, Amer. Math. Monthly **79** (1972), 571–586; correction, ibid. 80 (1973), 281. MR 0308084 (46 #7199)