Lectures on Serre's conjectures

Kenneth A. Ribet
William A. Stein

Contents

IAS/Park City Mathematics Series
Volume 00, November 23, 2010

Lectures on Serre's conjectures[1]

Kenneth A. Ribet
William A. Stein

[1]Math Department, MC 3840; Berkeley, CA 94720-3840.
E-mail address: ribet@math.berkeley.edu, was@math.berkeley.edu.

1991 *Mathematics Subject Classification.* 11

Key words and phrases. Serre's conjectures, modular forms, Galois representations

The second author was supported by a Cal@SiliconValley fellowship.

Preface

We shall begin by discussing some examples of mod ℓ representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. We'll try to motivate Serre's conjectures by referring first to the case of representations that are unramified outside ℓ; these should come from cusp forms on the full modular group $\mathrm{SL}(2, \mathbf{Z})$. In another direction, one might think about representations coming from ℓ-division points on elliptic curves, or more generally from ℓ-division points on abelian varieties of "GL_2-type." Amazingly, Serre's conjectures imply that all odd irreducible two-dimensional mod ℓ representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ may be realized in spaces of ℓ-division points on such abelian varieties. The weak Serre conjecture states that all such representations come from modular forms, and then it takes only a bit of technique to show that one can take the modular forms to have weight two (if one allows powers of ℓ in the level).

Since little work has been done toward proving the weak Serre conjecture, these notes will focus on the bridge between the weak and the strong conjectures. The latter states that each ρ as above comes from the space of cusp forms of a specific weight and level, with these invariants between determined by the local behavior of ρ at ℓ and at primes other than ℓ (respectively). To motivate the strong conjecture, and to begin constructing the bridge, we discuss the local behavior of those ρ that do come from modular forms. For the most part, we look only at forms of weight $k \geq 2$ whose levels N are prime to ℓ. For these forms, the behavior of ρ at ℓ is described in detail in [**32**], where theorems of P. Deligne and Fontaine are recalled. (In [**32**, §6], B. Edixhoven presents a proof of Fontaine's theorem.) Further, the behavior of ρ at primes $p \neq \ell$ may be deduced from H. Carayol's theorems [**11, 12**], which relate the behavior at p of the ℓ-adic representations attached to f with the p-adic component of the automorphic representation of $\mathrm{GL}(2)$ that one associates with f. (The behavior of ρ at ℓ in the case where ℓ divides N is analyzed in [**89**].)

In [**102**], Serre associates to each ρ a level $N(\rho)$ and a weight $k(\rho)$. These invariants are defined so that $N(\rho)$ is prime to ℓ and so that $k(\rho)$ is an integer greater than 1. As Serre anticipated, if ρ arises from a modular form of weight k and level N, and if k is at least 2 and N is prime to ℓ, then one has $k(\rho) \leq k$ and $N(\rho) \mid N$. To find an f for which $N = N(\rho)$ and $k = k(\rho)$ is to "optimize" the level and weight of a form giving ρ. As Edixhoven explains in his article [**32**], weight optimization follows in a somewhat straightforward manner from the theorems of Deligne and Fontaine alluded to above, Tate's theory of θ-cycles, and Gross's theorem on companion forms [**46**] (see also [**17**]). Moreover, it is largely the case that weight and level optimization can be performed independently.

In [**12**], Carayol analyzes the level optimization problem. He shows, in particular, that the problem breaks down into a series of sub-problems, all but one of which he treats by appealing to a single lemma, the lemma of [**12**, §3]. The remaining sub-problem is the one that intervenes in establishing the implication "Shimura-Taniyama $\Longrightarrow$ Fermat." This problem has been discussed repeatedly [**83, 84, 86, 87**]. In Section 3.10, we will explain the principle of [**86**].

The case $\ell = 2$ is the only remaining case for which the level optimization problem has not been resolved. The proof in [**26, 87**] of level optimization for $\ell \geq 3$ does not fully exploit multiplicity one results, but appears to completely break down when $\ell = 2$. In the recent paper [**9**], Kevin Buzzard observed that

many new cases of multiplicity one are known and that this can be used to obtain new level optimization results when $\ell = 2$.

In view of these remarks it might be appropriate for us to summarize in a few sentence what is known about the implication "weak Serre conjecture $\Longrightarrow$ strong Serre conjecture." As explained in [**26**], for $\ell \geq 5$ the weak conjecture of Serre implies the strong conjecture about the optimal weight, level, and character. For $\ell = 3$, the weak conjecture implies the strong conjecture, except in a few well-understood situations, where the order of the character must be divisible by ℓ when the level is optimal. The difficulty disappears if one works instead with Katz's definition of a mod ℓ modular form, where the character is naturally defined only mod ℓ. The situation is less complete when $\ell = 2$, but quite favorable. When $\ell = 2$ the situation concerning the weight is explained by Edixhoven in [**32**]: the results of [**17**] do not apply and those of [**46**] rely on unchecked compatibilities.

A certain amount of work has been done on the Hilbert modular case, i.e., the case where $\mathbf{Q}$ is replaced by a totally real number field F. For this work, the reader may consult articles of Frazer Jarvis [**52, 53, 54**], Kazuhiro Fujiwara [**45**], and Ali Rajaei [**79**]. The authors are especially grateful to Fujiwara for sending them a preliminary version of his manuscript, "Level optimization in the totally real case." However, these notes will treat only the classical case $F = \mathbf{Q}$.

This paper emerged out of a series of lectures that were delivered by the first author at the 1999 IAS/Park City Mathematics Institute. The second author created the text based on the lectures and added examples, diagrams, an exercise section, and the index. Brian Conrad contributed the appendix, which describes a construction of Shimura.

For other expository accounts of Serre's conjectures, the reader may consult the articles of Edixhoven [**33, 34, 35**], H. Darmon [**22**], and R. Coleman [**15**].

The authors would like to thank K. Buzzard and Serre for many useful comments on various drafts of this paper, B. Conrad for providing the appendix, M. Emerton for his enlightening lecture on Katz's definition of modular forms, N. Jochnowitz for suggestions that improved the exposition in Section 2.2, and L. Kilford for help in finding examples of mod 2 representations in Section 3.7.

Kenneth A. Ribet
William A. Stein
University of California, Berkeley

CHAPTER 1

Introduction to Serre's conjecture

1.1. Introduction

Let's start with an elliptic curve $E/\mathbf{Q}$. Nowadays, it's a familiar activity to consider the Galois representations defined by groups of division points of E. Namely, let n be a positive integer, and let $E[n]$ be the kernel of multiplication by n on $E(\overline{\mathbf{Q}})$. The group $E[n]$ is free of rank two over $\mathbf{Z}/n\mathbf{Z}$. After a choice of basis, the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $E[n]$ is given by a homomorphism

$$\rho_n : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(E[n]) \approx \mathrm{GL}(2, \mathbf{Z}/n\mathbf{Z}).$$

This homomorphism is unramified at each prime p that is prime to the product of n with the conductor of E (see Exercise 15). For each such p, the element $\rho_n(\mathrm{Frob}_p)$ is a 2×2 matrix that is well defined up to conjugation. Its determinant is p mod n; its trace is a_p mod n, where a_p is the usual "trace of Frobenius" attached to E and p, i.e., the quantity $1 + p - \#E(\mathbf{F}_p)$. In his 1966 article [**107**], G. Shimura studied these representations and the number fields that they cut out for the curve $E = J_0(11)$. (This curve was also studied by Serre [**91**, pg. 254].) He noticed that for prime values $n = \ell$, the representations ρ_n tended to have large images. In [**93**], J-P. Serre proved that for any fixed elliptic curve E, not having complex multiplication, the indices

$$[\mathrm{GL}(2, \mathbf{Z}/n\mathbf{Z}) : \rho_n(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))]$$

are bounded independently of n. In Shimura's example, Serre proved that

$$[\mathrm{GL}(2, \mathbf{Z}/\ell\mathbf{Z}) : \rho_\ell(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))] = 1$$

for all $\ell \neq 5$ (see [**93**, §5.5.1]).

In this article, we will be concerned mainly with two-dimensional representations over finite fields. To that end, we restrict attention to the case where $n = \ell$ is prime. The representation ρ_ℓ is "modular" in the familiar sense that it's a representation of a group over a field in positive characteristic. The theme of this course is that it's modular in a different and deeper sense: it comes from a modular form! Indeed, according to a recent preprint of Breuil, Conrad, Diamond and Taylor (see [**7, 19, 114, 117**]), the Shimura-Taniyama conjecture is now a theorem—all elliptic curves over $\mathbf{Q}$ are modular!! Thus if N is the conductor of E, there is a weight-two newform $f = \sum_{n=1}^{\infty} c_n q^n$ $(q = e^{2\pi i z})$ on $\Gamma_0(N)$ with the property that $a_p = c_p$ for all p prime to N. Accordingly, ρ_ℓ is connected up with modular forms via the relation $\mathrm{tr}(\rho_\ell(\mathrm{Frob}_p)) \equiv c_p \pmod{\ell}$, valid for all but finitely many primes p.

The Shimura-Taniyama conjecture asserts that for each positive integer N there is a bijection between isogeny classes of elliptic curves A over $\mathbf{Q}$ of conductor N and rational newforms f on $\Gamma_0(N)$ of weight two. Given A, the Shimura-Taniyama

conjecture produces a modular form $f = \sum_{n=1}^{\infty} c_n q^n$ whose Dirichlet series is equal to the L-series of A:

$$\sum_{n=1}^{\infty} \frac{c_n}{n^s} = L(f,s) \stackrel{?}{=\!=\!=} L(A,s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \ .$$

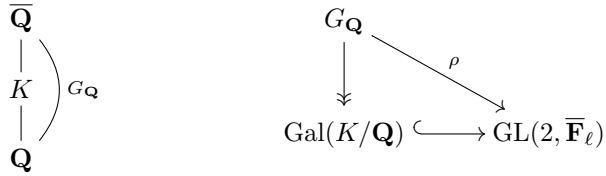
The integers a_n encode information about the number of points on A over various finite fields $\mathbf{F}_p$. If p is a prime not dividing N, then $a_p = p + 1 - \#A(\mathbf{F}_p)$; if $p \mid N$,

$$a_p = \begin{cases} -1 & \text{if } A \text{ has non-split multiplicative reduction at } p \\ 1 & \text{if } A \text{ has split multiplicative reduction at } p \\ 0 & \text{if } A \text{ has additive reduction at } p. \end{cases}$$

The integers a_n are obtained recursively from the a_p as follows:

- $a_{p^r} = \begin{cases} a_{p^{r-1}} a_p - p a_{p^{r-2}} & \text{if } p \nmid N \\ a_p^r & \text{if } p \mid N \end{cases}$
- $a_{nm} = a_n a_m, \qquad \text{if } (n,m) = 1.$

The conjectures made by Serre in [**102**], which are the subject of this paper, concern representations $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$. We always require (usually tacitly) that our representations are continuous. The continuity condition just means that the kernel of ρ is open, so that it corresponds to a finite Galois extension K of $\mathbf{Q}$. The representation ρ then embeds $\mathrm{Gal}(K/\mathbf{Q})$ into $\mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$. Since K is a finite extension of $\mathbf{Q}$, the image of ρ is contained in $\mathrm{GL}(2, \mathbf{F})$ for some finite subfield $\mathbf{F}$ of $\overline{\mathbf{F}}_\ell$.



For various technical reasons, the original conjectures of Serre insist that ρ be irreducible. It is nevertheless fruitful to consider the reducible case as well (see [**111**]).

The conjectures state (in particular) that each continuous irreducible ρ that satisfies a necessary parity condition "arises from" (or is associated with) a suitable modular form mod ℓ. To explain what's going on, let's start with

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

the unique (normalized) cusp form of weight 12 on $\mathrm{SL}(2, \mathbf{Z})$. In [**92**], Serre conjectured the existence of a "strictly compatible" family of ℓ-adic representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ whose L-function is the L-function of Δ, namely

$$L(\Delta, s) = \sum_{n=1}^{\infty} \tau(n) n^{-s} = \prod_{p} (1 - \tau(p) p^{-s} + p^{11-2s})^{-1},$$

where the product is taken over all prime numbers p. The conjectured ℓ-adic representations were constructed soon after by Deligne [**24**]. Specifically, Deligne constructed, for each prime ℓ, a representation

$$\rho_{\ell^\infty} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \mathbf{Z}_\ell),$$

unramified outside ℓ, such that for all primes $p \neq \ell$,

$$\mathrm{tr}(\rho_{\ell^\infty}(\mathrm{Frob}_p)) = \tau(p), \qquad \det(\rho_{\ell^\infty}(\mathrm{Frob}_p)) = p^{11}.$$

On reducing ρ_{ℓ^∞} mod ℓ, we obtain a representation

$$\rho_\ell : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \mathbf{F}_\ell)$$

with analogous properties. (Equalities are replaced by congruences mod ℓ.) In other words, the ρ_ℓ for Δ are just like the ρ_ℓ for an elliptic curve E, except that the integers a_p are replaced by the corresponding values of the τ-function. The determinant of ρ_ℓ is the 11th power of the mod ℓ cyclotomic character $\chi : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}_\ell^*$, i.e., the character giving the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the group of ℓth roots of unity in $\overline{\mathbf{Q}}$ (see Section 1.5).

More generally, take a weight $k \geq 12$ and suppose that $f = \sum_n c_n q^n$ is a nonzero weight-k cusp form for $\mathrm{SL}(2, \mathbf{Z})$ that satisfies $f|T_n = c_n f$ for all $n \geq 1$, T_n being the nth Hecke operator on the space of cusp forms of weight k for $\mathrm{SL}(2, \mathbf{Z})$ (see Section 1.5). Then the complex numbers c_n ($n \geq 1$) are algebraic integers. Moreover, the field $E := \mathbf{Q}(\ldots c_n \ldots)$ generated by the c_n is a totally real number field (of finite degree over $\mathbf{Q}$). Thus the c_n all lie in the integer ring $\mathcal{O}_E$ of E. For each ring homomorphism $\varphi : \mathcal{O}_E \to \overline{\mathbf{F}}_\ell$, one finds a representation

$$\rho = \rho_\varphi : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell),$$

unramified outside ℓ, such that

$$\mathrm{tr}(\rho(\mathrm{Frob}_p)) = \varphi(c_p), \qquad \det(\rho(\mathrm{Frob}_p)) = p^{k-1}$$

for all $p \neq \ell$. We have $\det \rho = \chi^{k-1}$. Of course, there is no guarantee that ρ is irreducible. We can (and do) suppose that ρ is semisimple by replacing it by its semisimplification. Then ρ is determined up to isomorphism by the displayed trace and determinant conditions; this follows from the Cebotarev density theorem and the Brauer-Nesbitt theorem [**21**], which states that semisimple representations are determined by their characteristic polynomials.

It is important to note that k is necessarily an even integer; otherwise the space $S_k(\mathrm{SL}(2, \mathbf{Z}))$ of weight-k cusp forms on $\mathrm{SL}(2, \mathbf{Z})$ is easily seen to be 0. Thus the determinant χ^{k-1} of ρ is an odd power of χ. In particular, $\det \rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}_\ell^*$ is unramified outside ℓ and takes the value -1 on complex conjugations $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. It's a nice exercise to check that, conversely, all continuous homomorphisms with these properties are odd powers of χ (see Exercise 1).

In the early 1970s, Serre conjectured that all homomorphisms that are "like ρ" come from cusp forms of some weight on $\mathrm{SL}(2, \mathbf{Z})$. Namely, let

$$\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$$

be a continuous, irreducible representation that is (1) unramified outside ℓ and (2) of odd determinant, in the sense that $\det \rho(c) = -1 \in \overline{\mathbf{F}}_\ell$ for complex conjugations $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. In a May, 1973 letter to Tate, Serre conjectured that ρ is of the form ρ_φ. This means that there is a weight $k \geq 12$, an eigenform $f \in S_k(\mathrm{SL}(2, \mathbf{Z}))$, and a homomorphism $\varphi : \mathcal{O}_E \to \overline{\mathbf{F}}_\ell$ (where $\mathcal{O}_E$ is the ring of integers of the field generated by the coefficients of f) so that ρ_φ and ρ are isomorphic.

To investigate Serre's conjecture, it is fruitful to consider the operation $\rho \mapsto \rho \otimes \chi$ on representations. This "twisting" operation preserves the set of representations that come from modular forms. Indeed, let $\theta = q\frac{d}{dq}$ be the classical differential

operator $\sum a_n q^n \mapsto \sum n a_n q^n$. According to Serre and Swinnerton-Dyer [**61, 94, 112**], if f is a mod ℓ form of weight k, then θf is a mod ℓ form of weight $k+\ell+1$. Then if ρ is associated to f, $\rho \otimes \chi$ is associated with θf. According to a result of Atkin, Serre and Tate (see [**97**, Th. 3] and Section 2.1), if ρ comes from an eigenform in some space $S_k(\mathrm{SL}(2,\mathbf{Z}))$, then a suitable twist $\rho \otimes \chi^i$ of f comes from a form of weight $\leq \ell+1$.

Serre's conjecture thus has the following consequence: each two-dimensional irreducible odd representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ over $\overline{\mathbf{F}}_\ell$ that is unramified outside ℓ has a twist (by a power of χ) coming from an eigenform on $\mathrm{SL}(2,\mathbf{Z})$ of weight at most $\ell+1$. In particular, suppose that $\ell < 11$. Then the spaces $S_k(\mathrm{SL}(2,\mathbf{Z}))$ with $k \leq \ell+1$ are all 0; as a result, they contain no nonzero eigenforms! The conjecture that all ρ are modular (of level 1) thus predicts that there are *no* representations of the type contemplated if ℓ is 2, 3, 5 or 7. In support of the conjecture, the non-existence statement was proved for $\ell = 2$ by J. Tate in a July, 1973 letter to Serre [**113**]. Soon after, Serre treated the case $\ell = 3$ by methods similar to those of Tate. (See [**113**, p. 155] for a discussion and a reference to a note in Serre's Œuvres.) Quite recently, Sharon Brueggeman considered the case $\ell = 5$; she proved that the conjectured result follows from the Generalized Riemann Hypothesis (see [**8**]). In another direction, Hyunsuk Moon generalized Tate's result and proved that there are only finitely many isomorphism classes of continuous semisimple Galois representations $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_4(\overline{\mathbf{F}}_2)$ unramified outside 2 such that field $K/\mathbf{Q}$ corresponding to the kernel of ρ is totally real (see [**76**]). Similar work in this direction has been carried out by Joshi [**58**], under additional local hypothesis.

Serre discussed his conjecture with Deligne, who pointed out a number of surprising consequences. In particular, suppose that one takes a ρ coming from an eigenform f' of some weight and of level $N > 1$. On general grounds, ρ has the right to be ramified at primes p dividing N as well as at the prime ℓ. Suppose that, by accident as it were, ρ turned out to be unramified at all primes $p \mid N$. Then the conjecture would predict the existence of a level-1 form f' (presumably of the same weight as f) whose mod ℓ representation was isomorphic to ρ. For example, if $N = \ell^\alpha$ is a power of ℓ, then the conjecture predicts that ρ arises from a form f' of level 1. How could one manufacture the f'?

The passage $f \rightsquigarrow f'$ comes under the rubric of "level optimization". When you take a representation ρ that comes from high level N, and it seems as though that representation comes from a lower level N', then to "optimize the level" is to cough up a form of level N' that gives ρ.

Deligne pointed out also that Serre's conjecture implies that representations ρ over $\overline{\mathbf{F}}_\ell$ are required to "lift" to λ-adic representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. In the recent articles [**80**] and [**81**], R. Ramakrishna used purely Galois cohomological techniques to prove results in this direction.

1.2. The weak conjecture of Serre

In the mid 1980s, Gerhard Frey began lecturing on a link between Fermat's Last Theorem and elliptic curves (see [**42, 43**]). (Earlier, Hellegouarch had also considered links between Fermat's Last Theorem and elliptic curves; see the MathSciNet review and Appendix of [**48**].) As is now well known, Frey proposed that if $a^\ell + b^\ell$ was a perfect ℓth power, then the elliptic curve $y^2 = x(x-a^\ell)(x+b^\ell)$ could be proved to be non-modular. Soon after, Serre pointed out that the non-modularity

contemplated by Frey would follow from suitable level-optimization results concerning modular forms [**101**]. To formulate such optimization results, Serre went back to the tentative conjecture that he had made 15 years earlier and decided to study representations $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ that are not necessarily unramified at ℓ. The results, of course, were the conjectures of [**102**].

An important consequence of these conjectures is the so-called "weak conjecture of Serre." As background, we recall that Hecke eigenforms on congruence subgroups of $\mathrm{SL}(2, \mathbf{Z})$ give rise to two-dimensional representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. If we set things up correctly, we get representations over $\overline{\mathbf{F}}_\ell$. More specifically, take integers $k \geq 2$ and $N \geq 1$; these are the weight and level, respectively. Let $f = \sum a_n q^n$ be a normalized Hecke eigenform in the space $S_k(\Gamma_1(N))$ of complex weight-k cusp forms on the subgroup $\Gamma_1(N)$ of $\mathrm{SL}(2, \mathbf{Z})$. Thus f is nonzero and it satisfies $f|T_n = a_n f$ for all $n \geq 1$. Further, there is a character $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \mathbf{C}^*$ so that $f|\langle d \rangle = \varepsilon(d) f$ for all $d \in (\mathbf{Z}/N\mathbf{Z})^*$, where $\langle d \rangle$ is the diamond-bracket operator. Again, let $\mathcal{O}$ be the ring of integers of the field $\mathbf{Q}(\ldots a_n \ldots)$ generated by the a_n; this field is a number field that is either totally real or a CM field. Consider a ring homomorphism $\varphi : \mathcal{O} \to \overline{\mathbf{F}}_\ell$ as before. Associated to this set-up is a representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ with properties that connect it up with f (and φ). First, the representation is unramified at all p not dividing ℓN. Next, for all such p, we have

$$\mathrm{tr}(\rho(\mathrm{Frob}_p)) = a_p, \qquad \det(\rho(\mathrm{Frob}_p)) = p^{k-1}\varepsilon(p);$$

the numbers a_p and $p^{k-1}\varepsilon(p)$, literally in $\mathcal{O}$, are mapped tacitly into $\overline{\mathbf{F}}_\ell$ by φ. The representation ρ is determined up to isomorphism by the trace and determinant identities that are displayed, plus the supplemental requirement that it be semisimple. We are interested mainly in the (generic) case in which ρ is irreducible; in that case, it is of course semisimple.

The construction of ρ from f, k and φ was described in [**24**]. In this article, Deligne sketches a method that manufactures for each non-archimedean prime λ of E a representation $\tilde{\rho}_\lambda : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, E_\lambda)$, where E_λ denotes the completion of E at λ. Given φ, we let $\lambda = \ker \varphi$ and find a model of $\tilde{\rho}_\lambda$ over the ring of integers $\mathcal{O}_\lambda$ of E_λ. Reducing $\tilde{\rho}_\lambda$ modulo λ, we obtain a representation over the finite field $\mathcal{O}_\lambda/\lambda\mathcal{O}_\lambda$, and φ embeds this field into $\overline{\mathbf{F}}_\ell$.

In fact, as Shimura has pointed out, the machinery of [**24**] can be avoided if one seeks only the mod λ representation attached to f (as opposed to the full λ-adic representation $\tilde{\rho}_\lambda$). As the first author pointed out in [**87**], one can use congruences among modular forms to find a form of weight two and level $N\ell^2$ that gives rise to ρ. Accordingly, one can find ρ concretely by looking within the group of ℓ-division points of a suitable abelian variety over $\mathbf{Q}$: the variety $J_1(\ell^2 N)$, which is defined in Section 2.3 and in Conrad's Appendix.

Which representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arise in this way (as k, N, f and φ all vary)? As in the case $N = 1$ (i.e., that where $\Gamma_1(N) = \mathrm{SL}(2, \mathbf{Z})$), any ρ that comes from modular forms is an odd representation: we have $\det(\rho(c)) = -1$ when $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is a complex conjugation. To see this, we begin with the fact that $\varepsilon(-1) = (-1)^k$, which generalizes (1.4); this follows from the functional equation that relates $f(\frac{az+b}{cz+d})$ to $f(z)$ when $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is an element of $\Gamma_0(N)$ (see Exercise 7). On the other hand, using the Cebotarev density theorem, we find that $\det \rho = \chi^{k-1}\varepsilon$, where χ is again the mod ℓ cyclotomic character and where ε is regarded now as a map $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_\ell^*$ in the obvious way, namely by composing $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{F}}_\ell^*$

with the mod N cyclotomic character. The value on c of the latter incarnation of ε is the number $\varepsilon(-1) = (-1)^k$. Since $\chi(c) = -1$, we deduce that $(\det \rho)(c) = -1$, as was claimed.

Serre's weak conjecture states that, conversely, every irreducible odd representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ is modular in the sense that it arises from some f and φ.

A concrete consequence of the conjecture is that all odd irreducible 2-dimensional Galois representations ρ come from abelian varieties over $\mathbf{Q}$. Given ρ, one should be able to find a totally real or CM number field E, an abelian variety A over $\mathbf{Q}$ of dimension $[E : \mathbf{Q}]$ that comes equipped with an action of the ring of integers $\mathcal{O}$ of E, and a ring homomorphism $\varphi : \mathcal{O} \to \overline{\mathbf{F}}_\ell$ with the following property: Let $\lambda = \ker \varphi$. *Then the representation* $A[\lambda] \otimes_{\mathcal{O}/\lambda} \overline{\mathbf{F}}_\ell$ *is isomorphic to* ρ. (In comparing $A[\lambda]$ and ρ, we use $\varphi : \mathcal{O}/\lambda \hookrightarrow \overline{\mathbf{F}}_\ell$ to promote the 2-dimensional $A[\lambda]$ into a representation over $\overline{\mathbf{F}}_\ell$.)

Much of the evidence for the weak conjecture concerns representations taking values in $\mathrm{GL}(2, \mathbf{F}_q)$ where the finite field $\mathbf{F}_q$ has small cardinality. In his original article [**102**, §5], Serre's discusses a large number of examples of such representations. Serre uses theorems of Langlands [**68**] and Tunnell [**115**] to establish his weak conjecture for odd irreducible representations with values in $\mathrm{GL}(2, \mathbf{F}_2)$ and $\mathrm{GL}(2, \mathbf{F}_3)$. Further, he reports on numerical computations of J.-F. Mestre that pertain to representations over $\mathbf{F}_4$ (and trivial determinant). Additionally, Serre remarks [**102**, p. 219] that the weak conjecture is true for those representations with values in $\mathrm{GL}(2, \overline{\mathbf{F}}_p)$ that are dihedral in the sense that their projective images (in $\mathrm{PGL}(2, \overline{\mathbf{F}}_p)$) are dihedral groups. (See also [**29**, §5] for a related argument.) This section of Serre's paper concludes with examples over $\mathbf{F}_9$ and $\mathbf{F}_7$.

More recently, representations over the fields $\mathbf{F}_4$ and $\mathbf{F}_5$ were treated, under somewhat mild hypotheses, by Shepherd-Barron and Taylor [**105**]. For example, Shepherd-Barron and Taylor show that $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \mathbf{F}_5)$ is isomorphic to the 5-torsion representaton of an elliptic curve over $\mathbf{Q}$ provided that $\det \rho$ is the mod 5 cyclotomic character. Because elliptic curves over $\mathbf{Q}$ are modular, it follows that ρ is modular.

1.3. The strong conjecture

Fix an odd irreducible Galois representation

$$\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell).$$

As discussed above, the weak conjecture asserts that ρ is modular, in the sense that there exists integers N and k such that ρ comes from some $f \in S_k(\Gamma_1(N))$. The *strong conjecture* goes further and gives a recipe for integers $N(\rho)$ and $k(\rho)$, then asserts that ρ comes from $S_{k(\rho)}(\Gamma_1(N(\rho)))$. In any particular instance, the strong conjecture is, a priori, easier to verify or disprove than the weak conjecture because $S_{k(\rho)}(\Gamma_1(N(\rho)))$ is a finite-dimensional vector space that can be computed (using, e.g., the algorithm in [**73**]). The relation between the weak and strong conjectures is analogous to the relation between the assertion that an elliptic curve is modular of some level and the assertion that an elliptic curve A is modular of a specific level, the conductor of A.

For each prime p, let $I_p \subset G_{\mathbf{Q}}$ denote an inertia group at p. The optimal level is a product

$$N(\rho) = \prod_{p \neq \ell} p^{n(p)},$$

where $n(p)$ depends only on $\rho|_{I_p}$. The optimal weight $k(\rho)$ depends only on $\rho|_{I_\ell}$. The integer $n(p)$ is a conductor in additive notation. In particular, $n(p) = 0$ if and only if ρ is unramified at p.

View ρ as a homomorphism $G_{\mathbf{Q}} \to \mathrm{Aut}(V)$, where V is a two-dimensional vector space over $\overline{\mathbf{F}}_\ell$. It is natural to consider the subspace of inertia invariants:

$$V^{I_p} := \{v \in V : \rho(\sigma)v = v, \text{ all } \sigma \in I_p\}.$$

For example, $V^{I_p} = V$ if and only if ρ is unramified at p. Define

$$n(p) := \dim(V/V^{I_p}) + \mathrm{Swan}(V),$$

where the wild term $\mathrm{Swan}(V)$ is the Swan conductor

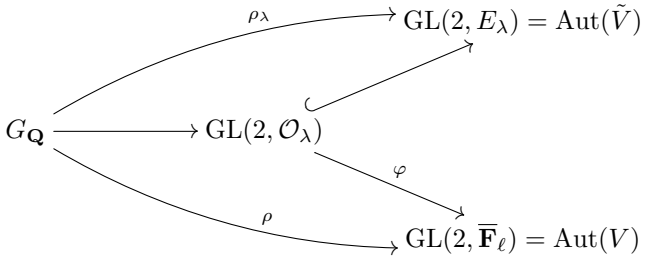
$$\mathrm{Swan}(V) := \sum_{i=1}^{\infty} \frac{1}{[G_0 : G_i]} \dim(V/V^{G_i}) \geq 0.$$

Here $G_0 = I_p$ and the $G_i \subset G_0$ are the higher ramification groups.

Suppose that ρ arises from a newform $f \in S_k(\Gamma_1(N))$. A theorem of Carayol [**12**], which was proved independently by Livné [**70**, Prop. 0.1], implies that $N(\rho) \mid N$. It is productive to study the quotient $N/N(\rho)$. Let $\mathcal{O}$ be the ring of integers of the field generated by the Fourier coefficients of f and let $\varphi : \mathcal{O} \to \overline{\mathbf{F}}_\ell$ be the map such that $\varphi(a_p) = \mathrm{tr}(\rho(\mathrm{Frob}_p))$. Let λ be a prime of $\mathcal{O}$ lying over ℓ and E_λ be the completion of $\mathrm{Frac}(\mathcal{O})$ at λ. Deligne [**24**] attached to the pair f, λ a representation

$$\rho_\lambda : G_{\mathbf{Q}} \to \mathrm{GL}(2, E_\lambda) = \mathrm{Aut}(\tilde{V})$$

where $\tilde{V}$ is a two-dimensional vector space over E_λ. The representation ρ_λ can be conjugated so that its images lies inside $\mathrm{GL}(2, \mathcal{O}_\lambda)$; the reduction of ρ_λ modulo λ is then ρ. The following diagram summarizes the set up:



Let $m(p)$ be the power of p dividing the conductor of ρ_λ. In [**12**], Carayol proves that $m(p) = \mathrm{ord}_p N$. As above, $m(p) = \dim(\tilde{V}/\tilde{V}^{I_p}) + (\text{wild term})$, and the wild term is the same as for ρ. The power of p dividing $N/N(\rho)$ is $\dim(\tilde{V}/\tilde{V}^{I_p}) - \dim(V/V^{I_p}) = \dim V^{I_p} - \dim \tilde{V}^{I_p}$. Though $\tilde{V}$ and V are vector spaces over different fields, we can compare the dimensions of their inertia invariant subspaces. The formula

$$\mathrm{ord}_p(N) = n(p) + (\dim V^{I_p} - \dim \tilde{V}^{I_p}) \tag{1.1}$$

indicates how this difference is the deviation of N from the optimal level locally at p. This is the description of $n(p)$ that is used in proving that if ρ is modular at

all, then it is possible to refine N and k to eventually discover that ρ arises from a newform in $S_{k(\rho)}(\Gamma_1(N(\rho)))$. After much work (see [**26, 87**]) it has been shown that *for $\ell > 2$ the weak and strong conjectures are equivalent.* See [**9**] for equivalence in many cases when $\ell = 2$.

Rearranging (1.1) into

$$n(p) = \mathrm{ord}_p(N) - (\dim V^{I_p} - \dim \tilde{V}^{I_p})$$

provides us with a way to read off $N(\rho)$ from $\mathrm{ord}_p(N)$, $\dim V^{I_p}$, and $\dim \tilde{V}^{I_p}$. If $f \in S_k(\Gamma_1(N))$ gives rise to ρ and $\ell \nmid N$, then $k(\rho) \leq k$. In contrast, if we allow powers of ℓ in the level then the weight k can always be made equal to 2.

1.4. Representations arising from an elliptic curve

Equations for elliptic curves can be found in the Antwerp tables [**4**] and the tables of Cremona [**20**]. For example, consider the elliptic curve B given by the equation $y^2+y = x^3+x^2-12x+2$. This is the curve labeled **141A1** in [**20**]; it has conductor $N = 3 \cdot 47$ and discriminant $\Delta = 3^7 \cdot 47$. There is a newform $f \in S_2(\Gamma_0(141))$ attached to B. Because N is square free, the elliptic curve B is *semistable*, in the sense that B has multiplicative reduction at each prime.

The curve B is isolated in its isogeny class; equivalently, for every ℓ the representation

$$\rho_\ell : G_{\mathbf{Q}} \to \mathrm{Aut}(B[\ell]) \approx \mathrm{GL}(2, \mathbf{F}_\ell)$$

is irreducible (see Exercise 4 and Exercise 5). We will frequently consider the representations ρ_ℓ attached to B. The following proposition shows that because B is semistable, each ρ_ℓ is surjective [**93**].

Proposition 1.1. *If A is a semistable elliptic curve over $\mathbf{Q}$ and ℓ is a prime such that ρ_ℓ is irreducible, then ρ_ℓ is surjective.*

Proof. Serre proved this when ℓ is odd; see [**93**, Prop. 21], [**103**, §3.1]. If ρ_2 isn't surjective, then by [**93**, Prop. 21(b)] and Theorem 2.10 it's unramified outside 2. This contradicts [**113**]. □

To give a flavor of Serre's invariants, we describe $N(\rho_\ell)$ and $k(\rho_\ell)$ for the representations ρ_ℓ attached to B. (Note that we still have not defined $k(\rho)$.) At primes p of bad reduction, after a possible unramified quadratic extension of $\mathbf{Q}_p$, the elliptic curve B is a Tate curve. This implies that for $p \neq \ell$, the representation ρ_ℓ is unramified at p if and only if $\mathrm{ord}_p(\Delta) \equiv 0 \pmod{\ell}$; for more details, see Section 2.4.

The optimal level $N(\rho_\ell)$ is a divisor of $3 \cdot 47$; it is divisible only by primes for which ρ_ℓ is ramified, and is not divisible by ℓ. The representation ρ_ℓ is unramified at 3 if and only if $\ell \mid \mathrm{ord}_3(\Delta) = 7$, i.e., when $\ell = 7$. Furthermore, ρ_ℓ is always ramified at 47. First suppose $\ell \notin \{3, 47\}$. If in addition $\ell \neq 7$ then $N(\rho_\ell) = 3 \cdot 47$, and $k(\rho_\ell) = 2$ since $\ell \nmid 3 \cdot 47$. If $\ell = 7$ then $N(\rho_\ell) = 47$, and again $k(\rho_\ell) = 2$. The remaining cases are $\ell = 3$ and $\ell = 47$. If $\ell = 47$ then $N(\rho_\ell) = 3$, and because $\ell - 1$ is the order of the cyclotomic character, $k(\rho_\ell) \equiv 2 \pmod{47-1}$; Serre's recipe then gives $k(\rho_\ell) = 2 + (47 - 1) = 48$. Similarly, when $\ell = 3$, we have $N(\rho_\ell) = 47$ and $k(\rho_\ell) = 2 + (3 - 1) = 4$. The following table summarizes the Serre invariants:

Table 1.4. The Serre invariants of ρ_ℓ

ℓ	$N(\rho_\ell)$	$k(\rho_\ell)$
3	47	4
7	47	2
47	3	48
all other ℓ	141	2

To verify the strong conjecture of Serre for $\ell = 3, 47$, we use a standard trade-off of level and weight, which relates eigenforms in $S_2(\Gamma_0(141); \mathbf{F}_\ell)$ to eigenforms in $S_{2+\ell-1}(\Gamma_0(141/\ell); \mathbf{F}_\ell)$ (see Section 3.1). The only exceptional prime is $\ell = 7$, for which the minimal weight $k(\rho)$ is 2. The strong conjecture of Serre predicts the existence of an eigenform $g \in S_2(\Gamma_0(47))$ that gives rise to ρ_ℓ. Our initial instinct is to look for an elliptic curve A of conductor 47 such that $A[\ell] \cong B[\ell]$, as $G_{\mathbf{Q}}$-modules. In fact, there are no elliptic curves of conductor 47. This is because $S_2(\Gamma_0(47))$ is four dimensional, having basis the Galois conjugates of a single eigenform $g = \sum c_n q^n$. The Fourier coefficients c_n of g generate the full ring of integers in the field K obtained from $\mathbf{Q}$ by adjoining a root of $h = x^4 - x^3 - 5x^2 + 5x - 1$. The discriminant $1957 = 19 \cdot 103$ of K equals the discriminant of h, so a root of h generates the full ring of integers. The eigenvalue c_2 satisfies $h(c_2) = 0$. Since $h \equiv (x+2)(x^3 + 4x^2 + x + 3) \pmod 7$, there is a prime λ lying over 7 such that $\mathcal{O}/\lambda \cong \mathbf{F}_7$; the isomorphism is given by $c_2 \mapsto -2 \bmod 7$. As a check, note that $\#B(\mathbf{F}_2) = 5$ so $a_2 = 3 - 5 = -2 = \varphi(c_2)$. More generally, for $p \nmid 7 \cdot 141$, we have $\varphi(c_p) \equiv a_p \bmod 7$. This equality of traces implies that the representation $\rho_{g,\lambda}$ is isomorphic to $\rho = \rho_{A,7}$, so A is modular of level 47.

1.5. Background material

In this section, we review the cyclotomic character, Frobenius elements, modular forms, and Tate curves. We frequently write $G_{\mathbf{Q}}$ for $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Many of these basics facts are also summarized in [**23**].

1.5.1. The cyclotomic character

The mod ℓ *cyclotomic character* is defined by considering the group $\boldsymbol{\mu}_\ell$ of ℓth roots of unity in $\overline{\mathbf{Q}}$; the action of the Galois group $G_{\mathbf{Q}}$ on the cyclic group $\boldsymbol{\mu}_\ell$ gives rise to a continuous homomorphism

$$\chi_\ell : G_{\mathbf{Q}} \to \mathrm{Aut}(\boldsymbol{\mu}_\ell). \tag{1.2}$$

Since $\boldsymbol{\mu}_\ell$ is a cyclic group of order ℓ, its group of automorphisms is canonically the group $(\mathbf{Z}/\ell\mathbf{Z})^* = \mathbf{F}_\ell^*$. We emerge with a map $G_{\mathbf{Q}} \to \mathbf{F}_\ell^*$, which is the character in question.

Let A be an elliptic curve and ℓ be a prime number. The Weil pairing e_ℓ (see [**109**, III.8]) sets up an isomorphism of $G_{\mathbf{Q}}$-modules

$$e_\ell : \bigwedge^2 A[\ell] \xrightarrow{\cong} \boldsymbol{\mu}_\ell. \tag{1.3}$$

The determinant of the representation $\rho_{A,\ell}$ is the mod ℓ cyclotomic character χ_ℓ.

Suppose now that $c \in G_{\mathbf{Q}}$ is the automorphism "complex conjugation." Then the determinant of $\rho_{A,\ell}(c)$ is $\chi_\ell(c)$. Now c operates on roots of unity by the map

$\zeta \mapsto \zeta^{-1}$, since roots of unity have absolute value 1. Accordingly,

$$\det \rho_{A,\ell}(c) = -1; \tag{1.4}$$

one says that $\rho_{A,\ell}$ is *odd*. If $\ell \neq 2$, then $\rho_{A,\ell}(c)$ is conjugate over $\overline{\mathbf{F}}_\ell$ to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ (Exercise 7). If $\ell = 2$ then the characteristic polynomial of $\rho_{A,\ell}(c)$ is $(x+1)^2$ so $\rho_{A,\ell}(c)$ is conjugate over $\overline{\mathbf{F}}_\ell$ to either the identity matrix or $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

1.5.2. Frobenius elements

Let K be a number field. The Galois group $\mathrm{Gal}(K/\mathbf{Q})$ leaves the ring $\mathcal{O}_K$ of integers of K invariant, so that one obtains an induced action on the ideals of $\mathcal{O}_K$. The set of prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$ lying over p (i.e., that contain p) is permuted under this action. For each $\mathfrak{p}$, the subgroup $D_\mathfrak{p}$ of $\mathrm{Gal}(K/\mathbf{Q})$ fixing $\mathfrak{p}$ is called the *decomposition group* of $\mathfrak{p}$. Meanwhile, $\mathbf{F}_\mathfrak{p} := \mathcal{O}_K/\mathfrak{p}$ is a finite extension of $\mathbf{F}_p$. The extension $\mathbf{F}_\mathfrak{p}/\mathbf{F}_p$ is necessarily Galois; its Galois group is cyclic, generated by the Frobenius automorphism $\varphi_p : x \mapsto x^p$ of $\mathbf{F}_\mathfrak{p}$. There is a natural surjective map $D_\mathfrak{p} \to \mathrm{Gal}(\mathbf{F}_\mathfrak{p}/\mathbf{F}_p)$; its injectivity is equivalent to the assertion that p is unramified in $K/\mathbf{Q}$. Therefore, whenever this assertion is true, there is a unique $\sigma_\mathfrak{p} \in D_\mathfrak{p}$ whose image in $\mathrm{Gal}(\mathbf{F}_\mathfrak{p}/\mathbf{F}_p)$ is φ_p. The automorphism $\sigma_\mathfrak{p}$ is then a well-defined element of $\mathrm{Gal}(K/\mathbf{Q})$, the Frobenius automorphism for $\mathfrak{p}$. The various $\mathfrak{p}$ are all conjugate under $\mathrm{Gal}(K/\mathbf{Q})$ and that the Frobenius automorphism for the conjugate of $\mathfrak{p}$ by g is $g\sigma_\mathfrak{p}g^{-1}$. In particular, the various $\sigma_\mathfrak{p}$ are all conjugate; this justifies the practice of writing σ_p for any one of them and stating that σ_p is well defined up to conjugation.

We next introduce the concept of Frobenius elements in $G_\mathbf{Q} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Let p again be a prime and let $\mathfrak{p}$ now be a prime of the ring of integers of $\overline{\mathbf{Q}}$ lying over p. To $\mathfrak{p}$ we associate: (1) its residue field $\mathbf{F}_\mathfrak{p}$, which is an algebraic closure of $\mathbf{F}_p$, and (2) a decomposition subgroup $D_\mathfrak{p}$ of $G_\mathbf{Q}$. There is again a surjective map $D_\mathfrak{p} \to \mathrm{Gal}(\mathbf{F}_\mathfrak{p}/\mathbf{F}_p)$. The Frobenius automorphism φ_p topologically generates the target group. We shall use the symbol Frob_p to denote any preimage of φ_p in any $D_\mathfrak{p}$ corresponding to a prime lying over p, and refer to Frob_p as a Frobenius element for p in $G_\mathbf{Q}$. This element is doubly ill defined. The ambiguity in Frob_p results from the circumstance that $\mathfrak{p}$ needs to be chosen and from the fact that $D_\mathfrak{p} \to \mathrm{Gal}(\mathbf{F}_\mathfrak{p}/\mathbf{F}_p)$ has a large kernel, the inertia subgroup $I_\mathfrak{p}$ of $D_\mathfrak{p}$. The usefulness of $\mathrm{Frob}_\mathfrak{p}$ stems from the fact that the various $\mathfrak{p}$ are all conjugate, so that likewise the subgroups $D_\mathfrak{p}$ and $I_\mathfrak{p}$ are conjugate. Thus if ρ is a homomorphism mapping $G_\mathbf{Q}$ to some other group, the kernel of ρ contains one $I_\mathfrak{p}$ if and only if it contains all $I_\mathfrak{p}$. In this case, one says that ρ is *unramified* at p; the image of Frob_p is then an element of the target that is well defined up to conjugation.

Consider an elliptic curve A over $\mathbf{Q}$ and let ℓ be a prime number. The fixed field of $\rho_{A,\ell}$ is a finite Galois extension $K_\ell/\mathbf{Q}$ whose Galois group G_ℓ is a subgroup of $\mathrm{GL}(2, \mathbf{F}_\ell)$. A key piece of information about the extension $K_\ell/\mathbf{Q}$ is that its discriminant is divisible at most by ℓ and primes dividing the conductor of A. In other words, if $p \neq \ell$ is a prime number at which A has good reduction, then $K_\ell/\mathbf{Q}$ is unramified at ℓ (see Exercise 15); one says that the representation $\rho_{A,\ell}$ is unramified at p. Whenever this occurs, the construction described above produces a Frobenius element σ_p in G_ℓ that is well defined up to conjugation.

Fix again an elliptic curve A and a prime number ℓ, and let $\rho_{A,\ell} : G_\mathbf{Q} \to \mathrm{GL}(2, \mathbf{F}_\ell)$ be the associated representation. For each prime p not dividing ℓ at which A has good reduction the Frobenius $\sigma_p = \rho_{A,\ell}(\mathrm{Frob}_p)$ is well defined only up

to conjugation. Nevertheless, the trace and determinant of σ_p are well defined. The determinant of $\rho_{A,\ell}$ is the mod ℓ cyclotomic character χ, so $\sigma_p = \chi(\mathrm{Frob}_p) = p \in \mathbf{F}_\ell$. On the other hand, one has the striking congruence

$$\mathrm{tr}(\rho_{A,\ell}(\mathrm{Frob}_p)) = p + 1 - \#\tilde{A}(\mathbf{F}_p) \pmod{\ell}.$$

1.5.3. Modular forms

We now summarize some background material concerning modular forms. Serre's book [**96**] is an excellent introduction (it treats only $N = 1$). One might also read the survey article [**27**] or consult any of the standard references [**65, 66, 75, 108**].

The *modular group* $\mathrm{SL}(2, \mathbf{Z})$ is the group of 2×2 invertible integer matrices. For each positive integer N, consider the subgroup

$$\Gamma_1(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbf{Z}) \ : \ N \mid c \quad \text{and} \quad a \equiv d \equiv 1 \pmod{N} \right\}.$$

Let $\mathfrak{h}$ be the complex upper half plane. A *cusp form* of integer weight $k \geq 1$ and level N is a holomorphic function $f(z)$ on $\mathfrak{h}$ such that

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(N); \tag{1.5}$$

we also require that $f(z)$ vanishes at the cusps (see [**108**, §2.1]). We denote by $S_k(\Gamma_1(N))$ the space of weight-k cusp forms of level N. It is a finite dimensional complex vector space. When $k \geq 2$ a formula for the dimension can be found in [**108**, §2.6].

Modular forms are usually presented as convergent Fourier series

$$f(z) = \sum_{n=1}^{\infty} a_n q^n$$

where $q := e^{2\pi i z}$. This is possible because the matrices $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$ lie in $\Gamma_1(N)$ so that $f(z+b) = f(z)$ for all integers b. For the forms that most interest us, the complex numbers a_n are algebraic integers.

The space $S_k(\Gamma_1(N))$ is equipped with an action of $(\mathbf{Z}/N\mathbf{Z})^*$; this action is given by

$$f(z) \mapsto f|\langle \overline{d} \rangle(z) := (cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right)$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2, \mathbf{Z})$ is any matrix such that $d \equiv \overline{d} \pmod{N}$. The operator $\langle d \rangle = \langle \overline{d} \rangle$ is referred to as a "diamond-bracket" operator.

For each integer $n \geq 1$, the nth *Hecke operator* on $S_k(\Gamma_1(N))$ is an endomorphism T_n of $S_k(\Gamma_1(N))$. The action is generally written on the right: $f \mapsto f|T_n$. The various T_n commute with each other and are interrelated by identities that express a given T_n in terms of the Hecke operators indexed by the prime factors of n. If $p \nmid N$ is a prime define the operator T_p on $S_k(\Gamma_1(N))$ by

$$f|T_p(z) = \sum_{n=1}^{\infty} a_{np} q^n + p^{k-1} \sum_{n=1}^{\infty} a_n (f|\langle p \rangle) q^{np}.$$

For $p \mid N$ prime, define T_p by

$$f|T_p(z) = \sum_{n=1}^{\infty} a_{np} q^n.$$

The *Hecke algebra* associated to cusp forms of weight k on $\Gamma_1(N)$ is the subring

$$\mathbf{T} := \mathbf{Z}[\ldots T_n \ldots \langle d \rangle \ldots] \subset \mathrm{End}(S_k(\Gamma_1(N)))$$

generated by all of the T_n and $\langle d \rangle$. It is finite as a module over $\mathbf{Z}$ (see Exercise 20). The diamond-bracket operators are really Hecke operators, in the sense that they lie in the ring generated by the T_n; thus $\mathbf{T} = \mathbf{Z}[\ldots T_n \ldots]$.

An *eigenform* is a nonzero element $f \in S_k(\Gamma_1(N))$ that is a simultaneous eigenvector for every element of the Hecke algebra $\mathbf{T}$. Writing $f = \sum a_n q^n$ we find that $a_n = a_1 c_n$ where c_n is the eigenvalue of T_n on f. Since f is nonzero, a_1 is also nonzero, so it is possible to multiply f by a_1^{-1}. The resulting *normalized eigenform* wears its eigenvalues on its sleeve: $f = \sum c_n q^n$. Because f is an eigenform, the action of the diamond bracket operators defines a character $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \mathbf{C}^*$; we call ε the *character* of f.

Associated to an eigenform $f \in S_k(\Gamma_1(N))$ we have a system $(\ldots a_p \ldots)$, $p \nmid N$, of eigenvalues. We say that f is a *newform* if this system of eigenvalues is not the system of eigenvalues associated to an eigenform $g \in S_k(\Gamma_1(M))$ for some level $M \mid N$ with $M \neq N$. Newforms have been extensively studied (see [**2, 13, 69, 75**]); the idea is to understand where systems of eigenvalues first arise, and then reconstruct the full space $S_k(\Gamma_1(N))$ from newforms of various levels.

1.5.4. Tate curves

The Tate curve is a p-adic analogue of the exponentiation of the representation $\mathbf{C}/\Lambda$ of the group of an elliptic curve over $\mathbf{C}$. In this section we recall a few facts about Tate curves; for further details, see [**110**, V.3].

Let K be a finite extension of $\mathbf{Q}_p$; consider an elliptic curve E/K with *split multiplicative* reduction, and let j denote the j-invariant of E. By formally inverting the well-known relation

$$j(q(z)) = \frac{1}{q(z)} + 744 + 196884q(z) + \cdots$$

between the complex functions $q(z) = e^{2\pi i z}$ and $j(z)$, we find an element $q \in K^*$ with $j = j(q)$ and $|q| < 1$. There is a $\mathrm{Gal}(\overline{\mathbf{Q}}_p/K)$-equivariant isomorphism $E(\overline{\mathbf{Q}}_p) \cong \overline{\mathbf{Q}}_p^*/q^{\mathbf{Z}}$. The Tate curve, which we suggestively denote by $\mathbf{G}_m/q^{\mathbf{Z}}$, is a scheme whose $\overline{\mathbf{Q}}_p$ points equal $\overline{\mathbf{Q}}_p^*/q^{\mathbf{Z}}$.

As a consequence, the group of n-torsion points on the Tate curve is identified with the $\mathrm{Gal}(\overline{\mathbf{Q}}_p/K)$-module $\{\zeta_n^a (q^{1/n})^b : 0 \leq a, b \leq n-1\}$; here ζ_n is a primitive nth root of unity and $q^{1/n}$ is a fixed nth root of q in $\overline{\mathbf{Q}}_p$. In particular, the subgroup generated by ζ_n is invariant under $\mathrm{Gal}(\overline{\mathbf{Q}}_p/K)$, so the local Galois representation on $E[n]$ is reducible. It is also known that the group of connected component of the reduction of the Néron model of E over $\overline{\mathbf{F}}_p$ is a cyclic group whose order is $\mathrm{ord}_p(q)$. The situation is summarized by the following table (taken from [**88**]):

Complex case	p-adic case
$\mathbf{C}/\Lambda$	no p-adic analogue
$\downarrow$ exponential map $e^{2\pi iz}$	no exponential available
$\mathbf{C}^*/q^{\mathbf{Z}}$	$K^*/q^{\mathbf{Z}}$.

Remark 1.2. When E has non-split multiplicative reduction over K, there is an unramified extension L over which E aquires split multiplicative reduction.

1.5.5. Mod ℓ modular forms

There are several excellent papers to consult when learning about mod ℓ modular forms. The papers of Serre [**95**] and Swinnerton-Dyer [**112**] approach the subject from the point of view of Galois representations. Katz's paper [**59**] is very geometric. Edixhoven's paper [**32**] contains a clear description of the basic facts. See also Jochnowitz's paper [**56**].

CHAPTER 2

Optimizing the weight

In [**102**, §2] Serre associated to an odd irreducible Galois representation

$$\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$$

two integers $N(\rho)$ and $k(\rho)$, which are meant to be the minimal level and weight of a form giving rise to ρ.

Conjecture 2.1 (Strong conjecture of Serre)**.** Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be an odd irreducible Galois representation arising from a modular form. Then ρ arises from a modular form of level $N(\rho)$ and weight $k(\rho)$.

In this chapter, we are concerned with $k(\rho)$. We consider a mod ℓ representation ρ that arises from an eigenform of level N not divisible by ℓ. Using results of Fontaine and Deligne, we motivate Serre's recipe for $k(\rho)$. In [**32**], Edixhoven also defines an "optimal" weight, which sometimes differs from Serre's $k(\rho)$. Our definition is an "average" of the two; for example, we introduce a tiny modification of $k(\rho)$ when $\ell = 2$. We appologize for any confusion this may cause the reader.

Using various arguments involving the Eichler-Shimura correspondence and Tate's θ-cycles, Edixhoven showed in [**31**] that there must exist another form of weight at most $k(\rho)$, also of level N, which gives rise to ρ. Some of Edixhoven's result rely on unchecked compatibilities that are assumed in [**46**]; however, when $\ell \neq 2$ these results were obtained unconditionally by Coleman and Voloch in [**17**]. We sketch some of Edixhoven's arguments to convey the flavor of the subject.

Remark 2.2 (Notation)**.** We pause to describe a notational shorthand which we will employ extensively in this chapter. If $\rho : G \to \mathrm{Aut}(V)$ is a two-dimensional representation over a field $\mathbf{F}$, we will frequently write

$$\rho \sim \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

to mean that there is a basis for V with respect to which

$$\rho(x) = \begin{pmatrix} \alpha(x) & \beta(x) \\ \gamma(x) & \delta(x) \end{pmatrix} \in \mathrm{GL}_2(\mathbf{F})$$

for all $x \in G$. If we do not wish to specify one of the entries we will simply write $*$. Thus "$\rho \sim \left(\begin{smallmatrix} * & * \\ 0 & 1 \end{smallmatrix}\right)$" means that ρ possesses a one-dimensional invariant subspace, and the action on the quotient is trivial.

2.1. Representations arising from forms of low weight

We first consider irreducible Galois representations associated to newforms of low weight. Fix a prime ℓ and suppose $f = \sum a_n q^n$ is a newform of weight k and

level N, such that $\ell \nmid N$ and $2 \leq k \leq \ell + 1$. Let $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \mathbf{C}^*$ denote the character of f. Fix a homomorphism φ from the ring of integers $\mathcal{O}$ of $\mathbf{Q}(\ldots a_n \ldots)$ to $\overline{\mathbf{F}}_\ell$. To abbreviate, we often write a_n for $\varphi(a_n)$; thereby thinking of a_n as an element $\overline{\mathbf{F}}_\ell$. Let $\rho = \rho_{f,\varphi} : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be the two-dimensional semisimple odd Galois representation attached to f and φ, and assume that ρ is irreducible.

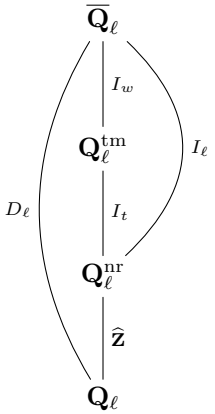
The recipe for $N(\rho)$ depends on the local behavior of ρ at primes p other than ℓ; the recipe for $k(\rho)$ depends on the restriction $\rho|_{I_\ell}$ of ρ to the inertia group at ℓ. Motivated by questions of Serre, Fontaine and Deligne described $\rho|_{I_\ell}$ in many situations. We distinguish two cases: the ordinary case and the non-ordinary case, which we call the *"supersingular case."*

2.1.1. The ordinary case

Deligne (see [**46**, Prop. 12.1]) considered the *ordinary case*, in which ρ arises from a weight-k newform f with $a_\ell(f) \neq 0 \in \overline{\mathbf{F}}_\ell$. He showed that ρ has a one-dimensional unramified quotient β, so $\rho|_{D_\ell} \sim \begin{pmatrix} \alpha & * \\ 0 & \beta \end{pmatrix}$ with $\beta(I_\ell) = 1$ and $\alpha\beta = \chi^{k-1}\varepsilon$. The mod N character ε is also unramified at ℓ because $\ell \nmid N$. Since the mod ℓ cyclotomic character χ has order $\ell - 1$ and $\rho|_{I_\ell} \sim \begin{pmatrix} \chi^{k-1} & * \\ 0 & 1 \end{pmatrix}$, the value of k modulo $\ell - 1$ is determined by $\rho|_{I_\ell}$. In the case when k is not congruent to 2 modulo $\ell - 1$, the restriction $\rho|_{I_\ell}$ determines the minimal weight $k(\rho)$. We will discuss the remaining case in Section 2.2.

2.1.2. The supersingular case and fundamental characters

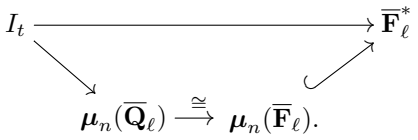
Fontaine (see [**32**, §6]) investigated the supersingular case, in which ρ arises from a newform f with $a_\ell(f) = 0 \in \overline{\mathbf{F}}_\ell$. We call such a newform f *supersingular*. To describe the restriction $\rho|_{I_\ell}$ of ρ to the inertia group at ℓ, we introduce the fundamental characters of the tame inertia group. Fix an algebraic closure $\overline{\mathbf{Q}}_\ell$ of the field $\mathbf{Q}_\ell$ of ℓ-adic numbers; let $\mathbf{Q}_\ell^{\mathrm{nr}} \subset \overline{\mathbf{Q}}_\ell$ denote the maximal unramified extension of $\mathbf{Q}_\ell$, and $\mathbf{Q}_\ell^{\mathrm{tm}} \subset \overline{\mathbf{Q}}_\ell$ the maximal tamely ramified extension of $\mathbf{Q}_\ell^{\mathrm{nr}}$. The extension $\mathbf{Q}_\ell^{\mathrm{tm}}$ is the compositum of all finite extensions of $\mathbf{Q}_\ell^{\mathrm{nr}}$ in $\overline{\mathbf{Q}}_\ell$ of degree prime to ℓ. Letting D_ℓ denote the decomposition group, I_ℓ the inertia group, I_t the tame inertia group, and I_w the wild inertia group, we have the following diagram:



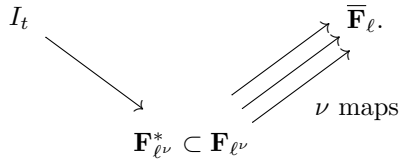
It is a standard fact (see, e.g., [**44**, §8]) that the extensions $\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})$, for all n not divisible by ℓ, generate $\mathbf{Q}_\ell^{\mathrm{tm}}$. For n not divisible by ℓ, the nth roots of unity $\boldsymbol{\mu}_n$ are contained in $\mathbf{Q}_\ell^{\mathrm{nr}}$. Kummer theory (see [**3**]) gives, for each n, a canonical isomorphism

$$\mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})/\mathbf{Q}_\ell^{\mathrm{nr}}) \xrightarrow{\sim} \boldsymbol{\mu}_n, \qquad \sigma \mapsto \frac{\sigma(\sqrt[n]{\ell})}{\sqrt[n]{\ell}}.$$

Each isomorphism lifts to a map $I_\ell \to \boldsymbol{\mu}_n$ that factors through the tame quotient I_t of I_ℓ. The groups $\boldsymbol{\mu}_n = \boldsymbol{\mu}_n(\overline{\mathbf{Q}}_\ell)$ lie in the ring of integers $\overline{\mathbf{Z}}_\ell$ of $\overline{\mathbf{Q}}_\ell$. Composing any of the maps $I_t \to \boldsymbol{\mu}_n$ with reduction modulo the maximal ideal of $\overline{\mathbf{Z}}_\ell$ gives a mod ℓ character $I_t \to \overline{\mathbf{F}}_\ell^*$, as illustrated:



Let $n = \ell^\nu - 1$ with $\nu > 0$. The map $I_t \to \boldsymbol{\mu}_n$ defines a character $\varepsilon : I_\ell \to \mathbf{F}_{\ell^\nu}^*$. Composing with each of the ν field embeddings $\mathbf{F}_{\ell^\nu} \to \overline{\mathbf{F}}_\ell$ gives the ν *fundamental characters* of level ν:



For example, the unique fundamental character of level 1 is the mod ℓ cyclotomic character (see Exercise 16). When $\nu = 2$, there are two fundamental characters, denoted Ψ and Ψ'; these satisfy $\Psi^\ell = \Psi'$ and $(\Psi')^\ell = \Psi$.

Let A be an elliptic curve over $\mathbf{Q}_\ell$ with good supersingular reduction. In [**93**], Serre proved that the representation

$$I_t \to \mathrm{Aut}(A[\ell]) \subset \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$$

is the direct sum of the two fundamental characters Ψ and Ψ'. One of the characters is

$$I_t \to \mathbf{F}_{\ell^2}^* \subset \mathrm{GL}(2, \mathbf{F}_\ell)$$

where $\mathbf{F}_{\ell^2}^*$ is contained in $\mathrm{GL}(2, \mathbf{F}_\ell)$ as a non-split Cartan subgroup of $\mathrm{GL}(2, \mathbf{F}_\ell)$. More precisely, $\mathbf{F}_{\ell^2}^*$ is embedded in $\mathrm{GL}(2, \mathbf{F}_\ell)$ via the action of the multiplicative group of a field on itself after a choice of basis. More generally, in unpublished joint work, Fontaine and Serre proved in 1979 that if f is a supersingular eigenform of weight $k \leq \ell$, then $\rho|_{I_\ell} : I_\ell \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ factors through I_t and is a direct sum of the two character Ψ^{k-1} and $(\Psi')^{k-1}$. Note that k is determined by this representation, because it is determined modulo $\ell^2 - 1$.

2.2. Representations of high weight

Let D_ℓ be a decomposition group at ℓ and consider a representation $\rho : D_\ell \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ that arises from a newform f of possibly large weight k. Let ρ^{ss} denote the *semisimplification* of ρ; so $\rho^{\mathrm{ss}} = \rho$ if ρ is irreducible, otherwise ρ^{ss} is a direct

sum of two characters α and β. The following lemma of Serre (see [**93**, Prop. 4]) asserts that ρ^{ss} is tamely ramified.

Lemma 2.3. *Any semisimple representation ρ is tame, in the sense that $\rho(I_w) = 0$.*

Proof. Since the direct sum of tame representations is tame, we may assume that ρ is simple.

The wild inertia group I_w is the profinite Sylow ℓ-subgroup of I_ℓ: it is a Sylow ℓ-subgroup because each finite Galois extension of $\mathbf{Q}_\ell^{\mathrm{tm}}$ has degree a power of ℓ, and the order of I_t is prime to ℓ; it is unique, because it is the kernel of $\mathrm{Gal}(\overline{\mathbf{Q}}_\ell/\mathbf{Q}_\ell) \to \mathrm{Gal}(\overline{\mathbf{Q}}_\ell/\mathbf{Q}_\ell^{\mathrm{tm}})$, hence normal.

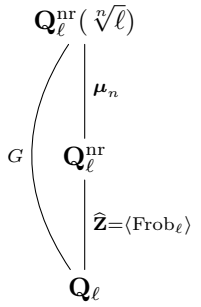
Because ρ is continuous, the image of D_ℓ is finite and we view ρ as a representation on a vector space W over a finite extension of $\mathbf{F}_\ell$. The subspace

$$W^{I_w} = \{w \in W : \sigma(\tau)w = w \text{ for all } \tau \in I_w\}$$

is invariant under D_ℓ. It is nonzero, as can be seen by writing the finite set W as a disjoint union of its orbits under I_w: since I_w is a pro-ℓ-group, each orbit has size either 1 or a positive power of ℓ. The orbit $\{0\}$ has size 1, and $\#W$ is a power of ℓ, so there must be at least $\ell - 1$ other singleton orbits $\{w\}$; for each of these, $w \in W^{I_w}$.

Since ρ is simple and W^{I_w} is a nonzero D_ℓ-submodule, it follows that $W^{I_w} = W$, as claimed. □

The restriction $\rho^{\mathrm{ss}}|_{I_\ell}$ is abelian and semisimple, so it is given by a pair of characters $\alpha, \beta : I_\ell \to \overline{\mathbf{F}}_\ell^*$. Let n be an integer not divisible by ℓ, and consider the tower of fields



in which $G = \mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})/\mathbf{Q}_\ell)$, $\boldsymbol{\mu}_n \cong \mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})/\mathbf{Q}_\ell^{\mathrm{nr}})$, and $\mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}/\mathbf{Q}_\ell)$ is topologically generated by a Frobenius element at ℓ. Choose a lift $g \in G$ of Frob_ℓ, and consider an element $h \in \boldsymbol{\mu}_n$ corresponding to an element $\sigma \in \mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})/\mathbf{Q}_\ell^{\mathrm{nr}})$. Then since g acts as the ℓth powering map on roots of unity,

$$\frac{g\sigma g^{-1}(\sqrt[n]{\ell})}{\sqrt[n]{\ell}} = \frac{g\sigma(\zeta_{g^{-1}}\sqrt[n]{\ell})}{\sqrt[n]{\ell}} = \frac{g(\zeta_{g^{-1}}h\sqrt[n]{\ell})}{\sqrt[n]{\ell}} = \frac{g(h)\sqrt[n]{\ell}}{\sqrt[n]{\ell}} = h^\ell.$$

Applying the conjugation formula $ghg^{-1} = h^\ell$ to ρ^{ss} gives $\rho^{\mathrm{ss}}(ghg^{-1}) = \rho^{\mathrm{ss}}(h^\ell) = \rho^{\mathrm{ss}}(h)^\ell$. The two representations $h \mapsto \rho^{\mathrm{ss}}(h)^\ell$ and $h \mapsto \rho^{\mathrm{ss}}(h)$ of I_t are thus equivalent via conjugation by $\rho^{\mathrm{ss}}(g)$; we have $\rho^{\mathrm{ss}}(g)\rho^{\mathrm{ss}}(h)\rho^{\mathrm{ss}}(g^{-1}) = \rho^{\mathrm{ss}}(ghg^{-1}) = \rho^{\mathrm{ss}}(h)^\ell$. Consequently, the pair of characters $\{\alpha, \beta\}$ is stable under the ℓth power map, so as a set $\{\alpha, \beta\} = \{\alpha^\ell, \beta^\ell\}$. There are two possibilities:

— The *ordinary case*: $\alpha^\ell = \alpha$ and $\beta^\ell = \beta$.

— The *supersingular case*: $\alpha^\ell = \beta \neq \alpha$ and $\beta^\ell = \alpha \neq \beta$.

In the first case α and β take values in $\mathbf{F}_\ell^*$ and in the second case they take values in $\mathbf{F}_{\ell^2}^*$ but not in $\mathbf{F}_\ell^*$. By the results discussed in Section 2.1, this terminology is consistent with the terminology introduced above.

We first consider the supersingular case. Let Ψ denote one of the fundamental characters of level 2, and write $\alpha = \Psi^n$, $\beta = \Psi^{n\ell}$, with n an integer modulo $\ell^2 - 1$. Next write the smallest non-negative representative for n in base ℓ: $n = a + \ell b$ with $0 \leq a, b \leq \ell - 1$. Then $\ell n \equiv b + \ell a \pmod{\ell^2 - 1}$. Switching α and β permutes a and b so, relabeling if necessary, we may assume that $a \leq b$. If $a = b$, then $\alpha = \Psi^a(\Psi')^a = \chi^a$, so α takes values in $\mathbf{F}_\ell^*$, which is not the supersingular case; thus we may assume that $0 \leq a < b \leq \ell - 1$. We now factor out by a power of the cyclotomic character:

$$\alpha = \Psi^n = \Psi^a(\Psi')^b = \Psi^a(\Psi')^a(\Psi')^{b-a} = \chi^a(\Psi')^{b-a}$$
$$\beta = \chi^a\Psi^{b-a}.$$

Put another way,

$$\rho^{\mathrm{ss}} \sim \chi^a \otimes \begin{pmatrix} \Psi^{b-a} & 0 \\ 0 & (\Psi')^{b-a} \end{pmatrix}.$$

The untwisted representation is $\left(\begin{smallmatrix} \Psi^{k-1} & 0 \\ 0 & (\Psi')^{k-1} \end{smallmatrix}\right)$, where $k = 1 + b - a$. Since $2 \leq 1+b-a \leq \ell-1$, the weight of the untwisted representation is in the range considered above. Thus we are in good shape to define $k(\rho)$.

Before giving $k(\rho)$ it is necessary to understand how the weight changes upon twisting by a power of the cyclotomic character χ. This problem is addressed by the theory of mod ℓ modular forms, first developed by Serre [**95**] and Swinnerton-Dyer [**112**], then generalized by Katz [**59**]. A brief review of the geometric theory, which gives an excellent definition of mod ℓ modular forms, can be found in [**32**, §2], [**35**, §1], or [**46**, §2].

In [**61**], Katz defined spaces of mod ℓ modular forms, and a q-expansion map

$$\alpha : \bigoplus_{k \geq 0} M_k(\Gamma_1(N); \mathbf{F}_\ell) \to \mathbf{F}_\ell[[q]].$$

This map is not injective, because both the Hasse invariant of weight $\ell - 1$ and the constant 1 have the same q-expansion.

Definition 2.4. The *minimal weight filtration* $w(f) \in \mathbf{Z}$ of an element f of the *ring* of mod ℓ modular forms is the smallest integer k such that the q-expansion of f comes from a modular form of weight k; if no such k exists, do not define $w(f)$.

Definition 2.5. Define the operator $\theta = q\frac{d}{dq}$ on q-expansions by $\theta(\sum a_n q^n) = \sum n a_n q^n$.

For example, if f is an eigenform of weight k, then there is a mod ℓ eigenform θf of weight $k + \ell + 1$, still of level N, whose q-expansion is $\theta(\sum a_n q^n)$.

Theorem 2.6. *Let f be a mod ℓ modular form. Then $w(\theta f) = w(f) + \ell + 1$ if and only if $\ell \nmid w(f)$. In addition, if $\ell \mid w(f)$ then $w(\theta f) < w(f) + \ell + 1$.*

2.2.1. The supersingular case

We now give Serre's recipe for $k(\rho)$ in the supersingular case. The minimal weight before twisting is $1+b-a$, which is a positive integer that is not divisible by ℓ. Each

twist by χ adds $\ell+1$ to the weight, so in the supersingular case we are motivated to define

$$k(\rho) := (1+b-a) + a(\ell+1) = 1 + \ell a + b.$$

We have to check that *at each step* the weight is prime to ℓ, so the minimal weight does not drop during any of the a twists by χ. Since $1 < 1+b-a < \ell$ and

$$(1+b-a) + a(\ell+1) \leq (\ell-1) + (\ell-2)(\ell+1) < \ell^2,$$

the weight can only drop if there exists c with $1 \leq c < a$ such that

$$(1+b-a) + c(\ell+1) \equiv 0 \pmod{\ell}.$$

If this occurred, then $c \equiv a-b-1 \pmod{\ell}$. But $1 \leq c < a \leq \ell-2$, so either $c = a-b-1$, which implies $c \leq 0$ since $a < b$, or $c = \ell+a-b-1 = a+\ell-1-b \geq a$, which would be a contradiction.

Assume that $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ arises from an eigenform f such that $a_\ell(f) = 0 \in \overline{\mathbf{F}}_\ell$. Now we sketch Edixhoven's proof that ρ arises from a mod ℓ eigenform of weight $k(\rho)$.

Let ρ^{ss} denote the semisimplification of the restriction of ρ to a decomposition group at ℓ. The restriction of ρ^{ss} to the inertia group at ℓ is

$$\rho^{\mathrm{ss}}|_{I_\ell} \sim \begin{pmatrix} \Psi^n & 0 \\ 0 & (\Psi')^n \end{pmatrix},$$

where Ψ and $\Psi' = \Psi^\ell$ are the two fundamental characters of level 2. If necessary, reorder Ψ and Ψ' so that $n = a + b\ell$ with $0 \leq a < b \leq \ell - 1$. Then

$$\Psi^n = \Psi^{a+b\ell} = \Psi^a(\Psi')^b = \Psi^a(\Psi')^a(\Psi')^{b-a} = \chi^a(\Psi')^{b-a},$$

and

$$\rho^{\mathrm{ss}}|_{I_\ell} \sim \chi^a \otimes \begin{pmatrix} (\Psi')^{b-a} & 0 \\ 0 & \Psi^{b-a} \end{pmatrix}.$$

Recall that, motivated by Fontaine's theorem on Galois representations arising from supersingular eigenforms, we defined

$$k(\rho) = a(\ell+1) + (b-a+1) = 1 + \ell a + b.$$

The first step in showing that ρ arises from a form of weight $k(\rho)$, is to recall the well known result that, up to twist, all systems of mod ℓ eigenvalues occur in weight at most $\ell+1$. This is the subject of the next section.

2.2.2. Systems of mod ℓ eigenvalues

Theorem 2.7. *Suppose ρ is modular of level N and some weight k, and that $\ell \nmid N$. Then some twist $\rho \otimes \chi^{-\alpha}$ is modular of weight $\leq \ell+1$ and level N.*

This is a general theorem, applying to both the ordinary and supersingular cases. See Serre [**97**, Th. 3] when $N = 1$; significant further work was carried out by Jochnowitz [**55**] and Ash-Stevens [**1**, Thm. 3.5] when $\ell \geq 5$. Two proofs are given in [**32**, Thm. 3.4 and §7]. The original method of Serre, Tate, and Koike for treating questions like this is to use the Eichler-Selberg trace formula. As Serre has pointed out to us, the weight appears in that formula simply as an exponent; this makes more or less clear that a congruence modulo $\ell^2 - 1$ gives information on modular forms mod ℓ.

As a digression, we pause to single out some of the tools involved in one possible proof of Theorem 2.7. Note that by twisting we may assume without loss of generality that $k \geq 2$. The group $\Gamma_1(N)$ acts by matrix multiplication on the real vector space $\mathbf{R}^2$. The Eichler-Shimura correspondence (see [**108**, §8.2]) is an isomorphism of real vector spaces

$$S_k(\Gamma_1(N)) \xrightarrow{\cong} H^1_P(\Gamma_1(N), \mathrm{Sym}^{k-2}(\mathbf{R}^2)).$$

The *parabolic* (or cuspidal) cohomology group H^1_P is the intersection, over all cusps $\alpha \in \mathbf{P}^1(\mathbf{Q})$, of the kernels of the restriction maps

$$\mathrm{res}_\alpha : H^1(\Gamma_1(N), \mathrm{Sym}^{k-2}(\mathbf{R}^2)) \to H^1(\Gamma_\alpha, \mathrm{Sym}^{k-2}(\mathbf{R}^2)),$$

where Γ_α denotes the stabilizer of α. For fixed z_0 in the upper half plane, the Eichler-Shimura isomorphism sends a cusp form f to the class of the cocycle $c : \Gamma_1(N) \to \mathrm{Sym}^{k-2}(\mathbf{R}^2)$ induced by

$$\gamma \mapsto \int_{z_0}^{\gamma(z_0)} \mathrm{Re}\left(f(z) \begin{pmatrix} z \\ 1 \end{pmatrix}^{k-2} dz \right),$$

where $\begin{pmatrix} z \\ 1 \end{pmatrix}^{k-2}$ denotes the image of $\begin{pmatrix} z \\ 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} z \\ 1 \end{pmatrix} \in \mathrm{Sym}^{k-2}(\mathbf{C}^2)$, and integration is coordinate wise. There is an action of the Hecke algebra $\mathbf{T}$ on

$$H^1_P(\Gamma_1(N), \mathrm{Sym}^{k-2}(\mathbf{R}^2)),$$

such that the Eichler-Shimura correspondence is an isomorphism of $\mathbf{T}$-modules.

The forms whose periods are integral form a lattice $H^1_P(\Gamma_1(N), \mathrm{Sym}^{k-2}(\mathbf{Z}^2))$ inside $H^1_P(\Gamma_1(N), \mathrm{Sym}^{k-2}(\mathbf{R}^2))$. Reducing this lattice modulo ℓ suggests that there is a relationship between mod ℓ modular forms and the cohomology group

$$H^1_P(\tilde{\Gamma}_1(N), \mathrm{Sym}^{k-2}(\mathbf{F}_\ell^2)),$$

where $\tilde{\Gamma}_1(N)$ is the image of $\Gamma_1(N)$ in $\mathrm{SL}(2, \mathbf{F}_\ell)$. Serre and Hida observed that for $k - 2 \geq \ell$ the $\tilde{\Gamma}_1(N)$ representations $\mathrm{Sym}^{k-2}(\mathbf{F}_\ell^2)$ are sums of representations arising in $\mathrm{Sym}^{k'-2}(\mathbf{F}_\ell^2)$ for $k' < k$. This essential idea is used in proving that all systems of eigenvalues occur in weight at most $\ell + 1$.

2.2.3. The supersingular case revisited

Let ρ be a supersingular mod ℓ representation that arises from some modular form. By Theorem 2.7 there is a form f of weight $k \leq \ell + 1$ such that $\chi^{-\alpha} \otimes \rho \sim \rho_f$. In fact, we may assume that $2 \leq k \leq \ell$; when $k = \ell + 1$ a theorem of Mazur (see [**32**, Thm. 2.8]) implies that there is a form of weight 2 giving rise to ρ_f, and when $k = 1$ we multiply f by the weight $\ell - 1$ Hasse invariant. To show that $w(\theta^\alpha f) = k(\rho)$ we investigate how application of the θ-operator changes the minimal weight. We have $(\rho_f \otimes \chi^\alpha)|_{I_\ell} \sim \begin{pmatrix} \Psi^n & 0 \\ 0 & (\Psi')^n \end{pmatrix}$ with $n = a + b\ell$ and $a < b$. Fontaine's theory (see Section 2.1) identifies the characters corresponding to $\rho_f|_{I_\ell}$ as powers Ψ^{k-1} and $(\Psi')^{k-1}$ of the fundamental characters. This gives an equality of unordered sets

$$\{\Psi^{k-1}\chi^\alpha, (\Psi')^{k-1}\chi^\alpha\} = \{\Psi^n, (\Psi')^n\}.$$

It is now possible to compute $w(\theta^\alpha f)$ by considering two cases, corresponding to the ways in which equality of unordered pairs can occur.

Case 1. Suppose that $\Psi^{k-1}\chi^\alpha = (\Psi')^n$. Since $\chi = \Psi^{\ell+1}$, we have

$$\Psi^{k-1+\alpha(\ell+1)} = \Psi^{k-1}\chi^\alpha = (\Psi')^n = (\Psi')^{a+b\ell} = \Psi^{b+a\ell}.$$

Comparing exponents of Ψ gives

$$k - 1 + \alpha(\ell+1) \equiv b + a\ell \pmod{\ell^2 - 1}, \tag{2.1}$$

which reduces modulo $\ell+1$ to $k-1 \equiv b-a \pmod{\ell+1}$; because $2 \le k \le \ell$, this implies that $k = 1+b-a$. Reducing (2.1) modulo $\ell-1$ and substituting $k = 1+b-a$ gives $b - a + 2\alpha \equiv b + a \pmod{\ell-1}$; we find the possible solutions $\alpha = a + m(\ell-1)/2$ with m an integer. No solution $\alpha = a + m(\ell-1)/2$, with m odd, satisfies (2.1), so $\alpha = a$ as an integer mod $\ell - 1$. Finally, we apply Theorem 2.6 and argue as in the end of Section 2.2, to show that

$$w(\theta^a f) = w(f) + a(\ell+1) = 1 + b - a + a\ell + a = 1 + b + a\ell = k(\rho).$$

Case 2. Suppose that $\Psi^{k-1}\chi^\alpha = \Psi^n$. Then

$$\Psi^{k-1+\alpha(\ell+1)} = \Psi^{k-1}\chi^\alpha = \Psi^n = \Psi^{a+b\ell}.$$

Comparing powers of Ψ, we obtain

$$k - 1 + \alpha(\ell+1) \equiv a + b\ell \pmod{\ell^2 - 1}, \tag{2.2}$$

which reduces modulo $\ell+1$ to $k-1 \equiv a-b \pmod{\ell+1}$; thus $k = \ell + 2 - (b-a)$. The difference $b-a$ must be greater than 1; otherwise $k = \ell+1$, contrary to our assumption that $2 \le k \le \ell$. Reducing (2.2) modulo $\ell-1$ gives

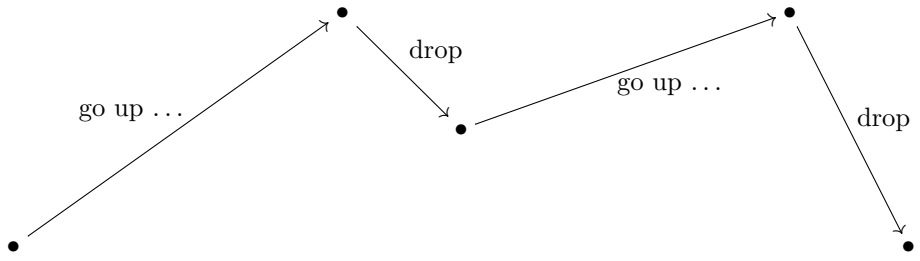
$$k - 1 + 2\alpha \equiv a + b \pmod{\ell-1};$$

substituting $k = \ell + 2 - (b-a)$ we find that $\alpha = b - 1 + m(\ell-1)/2$ with m an integer. If m is odd, then α does not satisfy (2.2), so $\alpha = b-1$ as an integer modulo $\ell - 1$. It remains to verify the equality $w(\theta^{b-1}f) = w(\rho)$. Unfortunately, $k = \ell + 2 - (b-a)$ is not especially telling. The argument of Case 1 does not apply to compute $w(\theta^\alpha f)$; instead we use θ-cycles.

Because f is supersingular, Fermat's Little Theorem implies that $\theta^{\ell-1}f = f$. We use Tate's theory of θ-cycles (see [**32**, §7] and [**55**]) to compute $w(\theta^{b-1}f)$. The θ-cycle associated to f is the sequence of integers

$$w(f), w(\theta f), w(\theta^2 f), \ldots, w(\theta^{\ell-2} f), w(f).$$

The θ-cycle for any supersingular eigenform must behave as follows (see Theorem 2.6):



go up ..., drop once, go up ..., drop to original weight

Knowing this, we can deduce the exact θ-cycle. List ℓ numbers starting and ending with k:

$$\begin{array}{l} k,\ k+(\ell+1),\ k+2(\ell+1),\ \ldots,\ k+(\ell-k)(\ell+1), \\ \ell+3-k,\ (\ell+3-k)+(\ell+1),\ \ldots,\ (\ell+3-k)+(k-3)(\ell+1), \\ k \end{array}$$

The first and second lines contain $\ell+1-k$ and $k-2$ numbers, respectively. All told, ℓ numbers are listed; this must be the θ-cycle.

It is now possible to compute $w(\theta^{b-1}f)$. If

$$b-1 \le \ell-k = \ell-(\ell+2-b+a) = -2+b-a,$$

then $a \le -1$, a contradiction; thus $b-1 > \ell-k$. It follows that

$$w(\theta^{b-1}f) = \ell+3-k+(\ell+1)(b-2-(\ell-k)) = 1+b+a\ell = k(\rho),$$

verifying Serre's conjecture in this case.

2.2.4. The ordinary case

We next turn to the ordinary case, in which

$$\rho|_{I_\ell} \sim \begin{pmatrix} \alpha & * \\ 0 & \beta \end{pmatrix}$$

with $\alpha, \beta : I_\ell \to \mathbf{F}_\ell^*$ powers of the cyclotomic character. View $\rho|_{I_\ell}$ as the twist of a representation in which the lower right entry is 1:

$$\begin{pmatrix} \alpha & * \\ 0 & \beta \end{pmatrix} \sim \beta \otimes \begin{pmatrix} \alpha\beta^{-1} & * \\ 0 & 1 \end{pmatrix}.$$

To determine the minimal weight of a form giving rise to $\rho|_{I_\ell}$, it is necessary to develop an ordinary version of θ-cycles. In general this is complicated, so we make the simplifying assumption that $\beta = 1$; then $\rho|_{I_\ell} \sim \begin{pmatrix} \chi^i & * \\ 0 & 1 \end{pmatrix}$ with $1 \le i \le \ell-1$. Deligne showed that if f is of weight k and $\beta = 1$, then the associated representation is $\left(\begin{smallmatrix} \chi^{k-1} & * \\ 0 & 1 \end{smallmatrix}\right)$ with $2 \le k \le \ell+1$. Motivated by this, our first reaction is to define $k(\rho)$ to be $i+1$. This definition does not distinguish between the extreme weights 2 and $\ell+1$ because they are congruent modulo $\ell-1$. Given a representation ρ arising from a form of weight either 2 or $\ell+1$, we cannot, in general, set $k(\rho) = 2$. For example, suppose $f = \Delta$ is the level 1 cusp form of weight 12 and ρ is the associated mod 11 representation. It would be wrong to set $k(\rho) = 2$, because there is no cusp form of weight 2 and level 1.

Warning: When $\ell = 2$ and our $k(\rho)$ is 3, Serre replaced $k(\rho)$ by 4 because there are no weight-3 modular forms whose character is of degree coprime to $\ell = 2$.

2.3. Distinguishing between weights 2 and $\ell+1$

We continue to motivate the definition of $k(\rho)$. Consider a representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ that arises from a newform f of the optimal level $N = N(\rho)$ and weight k satisfying $2 \le k \le \ell+1$. Assume that f is ordinary in the sense that $a_\ell(f) \ne 0 \in \overline{\mathbf{F}}_\ell$. Then, as discussed in Section 2.1,

$$\rho|_{I_\ell} \sim \left(\begin{smallmatrix} \chi^{k-1} & * \\ 0 & 1 \end{smallmatrix}\right),$$

so $\rho|_{I_\ell}$ determines k modulo $\ell - 1$. This suggests a way to define $k(\rho)$ purely in terms of the Galois representation ρ, at least when $k \notin \{2, \ell + 1\}$.

The key to defining $k(\rho)$ when $k = 2$ or $k = \ell + 1$ is good reduction. To understand why this is so, we briefly summarize Shimura's geometric construction of Galois representations associated to newforms of weight 2.

2.3.1. Geometric construction of Galois representations

Shimura attached mod ℓ representations to a weight-2 newform $f = \sum a_n q^n$ of level N. Let E be the totally real or CM field $\mathbf{Q}(\dots a_n \dots)$. In [**108**, Thm. 7.14], Shimura described how to associate to f an abelian variety $A = A_f$ over $\mathbf{Q}$ of dimension $[E : \mathbf{Q}]$ furnished with an embedding $E \hookrightarrow \mathrm{End}_{\mathbf{Q}} A$ (see also Conrad's appendix). The mod ℓ representations attached to f are then found in the ℓ-torsion of A.

Over the complex numbers, the abelian variety A is found as a quotient of the Jacobian of the Riemann surface

$$X_1(N) := \overline{\Gamma_1(N)\backslash\mathfrak{h}} = \Gamma_1(N)\backslash\mathfrak{h} \cup \{\text{cusps}\}.$$

The Riemann surface $X_1(N)$ has a structure of algebraic curve over $\mathbf{Q}$; it is called the *modular curve* of level N. Its Jacobian $J_1(N)$ is an abelian variety over $\mathbf{Q}$ which, by work of Igusa, has good reduction at all primes $\ell \nmid N$. The dimension of $J_1(N)$ equals the genus of $X_1(N)$; for example, when $N = 1$, the curve $X_1(1)$ is isomorphic over $\mathbf{Q}$ to the projective line and $J_1(1) = 0$. There are (at least) two functorial actions of the Hecke algebra $\mathbf{T}$ on $J_1(N)$, and (at least) two definitions of $J_1(N)$. In the next section we will fix choices, and then construct A as the quotient of $J_1(N)$ by the image of the annihilator in $\mathbf{T}$ of f.

2.3.1.1. *Hecke operators on $J_1(N)$.* We pause to formulate a careful definition of $X_1(N)$ and of our preferred functorial action of the Hecke operators T_p on $J_1(N)$. For simplicity, we assume that $N > 4$ and $p \nmid N$. Following [**46**, Prop. 2.1] there is a smooth, proper, geometrically connected algebraic curve $X_1(N)$ over $\mathbf{Z}[1/N]$ that represents the functor assigning to each $\mathbf{Z}[1/N]$-scheme S the set of isomorphism classes of pairs (E, α), where E is a generalized elliptic curve over S and $\alpha : (\boldsymbol{\mu}_N)_S \hookrightarrow E^{\mathrm{sm}}[N]$ an embedding of group schemes over S whose image meets every irreducible component in each geometric fiber. Let $X_1(N, p)$ be the fine moduli scheme over $\mathbf{Z}[1/N]$ that represents the functor assigning to each $\mathbf{Z}[1/N]$-scheme S the set of isomorphism classes of triples (E, α, C), where E is a generalized elliptic curve over S, $\alpha : (\boldsymbol{\mu}_N)_S \hookrightarrow E^{\mathrm{sm}}[N]$ an embedding of group schemes over S, and C a locally free subgroup scheme of rank p in $E^{\mathrm{sm}}[p]$, such that $\mathrm{im}(\alpha) \times C$ meets every irreducible component in each geometric fiber of E. Let $\pi_1, \pi_2 : X_1(N, p) \to X_1(N)$ over $\mathbf{Z}[1/N]$ be the two standard degeneracy maps, which are defined on genuine elliptic curves by $\pi_1(E, \alpha, C) = (E, \alpha)$ and $\pi_2(E, \alpha, C) = (E', \alpha' = \varphi\alpha)$, where $E' = E/C$ and $\varphi : E \to E'$ is the associated p-isogeny. The Hecke operator $T_p = (T_p)^*$ acts on divisors D on $X_1(N)_{/\mathbf{Q}}$ by

$$T_p(D) = (\pi_1)_* \circ \pi_2^* D.$$

For example, if (E, α) is a non-cuspidal $\overline{\mathbf{Q}}$-point, then

$$T_p(E, \alpha) = \sum (E', \varphi \circ \alpha \circ [p]^{-1}),$$

The Hecke operator $T_p = (T_p)^*$ acts on divisors D on $X_1(N)_{/\mathbf{Q}}$ by

$$T_p(D) = (\pi_1)_* \circ \pi_2^* D.$$

For example, if (E, α) is a non-cuspidal $\overline{\mathbf{Q}}$-point, then

$$T_p(E, \alpha) = \sum (E', \varphi \circ \alpha \circ [p]^{-1}),$$

where the sum is over all isogenies $\varphi : E \to E'$ of degree p, and $T_p = (T_p)^*$ acts on divisors D on $X_1(N)_{/\mathbf{Q}}$ by

$$T_p(D) = (\pi_1)_* \circ \pi_2^* D.$$

For example, if (E, α) is a non-cuspidal $\overline{\mathbf{Q}}$-point, then

$$T_p(E, \alpha) = \sum (E', \varphi \circ \alpha \circ [p]^{-1}),$$

where the sum is over all isogenies $\varphi : E \to E'$ of degree p, and where the sum is over all isogenies $\varphi : E \to E'$ of degree p, and $[p]^{-1}$ is the inverse of pth powering on $\boldsymbol{\mu}_N$. This map on divisors defines an endomorphism T_p of the Jacobian $J_1(N)$ associated to $X_1(N)$ via Picard functoriality.

For each prime p there is an involution $\langle p \rangle$ of $X_1(N)$ called a *diamond bracket operator*, defined functorially by

$$\langle p \rangle (E, \alpha) = (E, \alpha \circ [p]).$$

The diamond bracket operator defines a correspondence, such that the induced map $(\langle p \rangle)^*$ on $J_1(N)$ is

$$(\langle p \rangle)^* (E, \alpha) = (E, \alpha \circ [p^{-1}]).$$

If $(T_p)_*$ denotes the pth Hecke operator as defined in [**46**, §3], then

$$(T_p)_* = T_p \circ (\langle p^{-1} \rangle)^*,$$

Thus our T_p differs from Gross's $(T_p)_*$. Furthermore, upon embedding $X_1(N)$ into $J_1(N)$ and identifying weight-2 cusp forms with differentials on $J_1(N)$, Gross's $(T_p)_*$ induces, via Albanese functoriality, the usual Hecke action on cusp forms, whereas ours does not. In addition, we could have defined $X_1(N)$ by replacing the group scheme $\boldsymbol{\mu}_N$ by $(\mathbf{Z}/N\mathbf{Z})$. In this connection, see the discussion at the end of Section 5 of [**26**] and [**35**, §2.1].

2.3.1.2. *The representations attached to a newform.* Again let $\mathcal{O}$ be the ring of integers of $E = \mathbf{Q}(\ldots a_n \ldots)$, where $f = \sum a_n q^n$ is a weight-2 modular forms on $\Gamma_1(N)$. Recall that $A = A_f$ is the quotient of $J_1(N)$ by the image of the annihilator in $\mathbf{T}$ of f. In general, $\mathcal{O}$ need not be contained in $\operatorname{End} A$. However, by replacing A by an abelian variety $\mathbf{Q}$-isogenous to A, we may assume that $\mathcal{O}$ is contained in $\operatorname{End} A$ (see [**108**, pg. 199]). Let λ be a maximal ideal of $\mathcal{O}$ and set

$$A[\lambda] := \{P \in A(\overline{\mathbf{Q}}) : xP = 0 \text{ all } x \in \lambda\}.$$

By [**108**, Prop. 7.20, pg 190], $\dim_{\mathcal{O}/\lambda} A[\lambda] = 2$, so $A[\lambda]$ affords a 2-dimensional Galois representation, which is well-defined up to semisimplification. Let $\rho_{f,\lambda} : G_{\mathbf{Q}} \to A[\lambda]^{\mathrm{ss}}$ be the semisimplification of $A[\lambda]$.

2.3.1.3. *Good reduction.*

Definition 2.8. A finite group scheme G over $\mathbf{Q}_\ell^{\mathrm{nr}}$ is said to have *good reduction*, or to be *finite flat*, if it extends to a finite flat group scheme over the ring of integers $\mathcal{O}_{\mathbf{Q}_\ell^{\mathrm{nr}}}$ of $\mathbf{Q}_\ell^{\mathrm{nr}}$.

Proposition 2.9. *The representation $\rho_{f,\lambda}$ is finite flat at each prime $p \nmid N$.*

Proof. The finite flat group scheme extending $A[\lambda]$ is the scheme theoretic closure of $A[\lambda]$ in a good model $\mathcal{A}/\mathcal{O}_{\mathbf{Q}_\ell^{\mathrm{nr}}}$ of A. Such a model exists because A has good reduction at p. □

Consider again a Galois representation ρ as in the beginning of Section 2.3 such that $\rho|_{I_\ell} \sim \left(\begin{smallmatrix} \chi^{k-1} & * \\ 0 & 1 \end{smallmatrix}\right)$. If $k \not\equiv 2 \pmod{\ell-1}$ then $k(\rho)$ is defined to equal k. If $k \equiv 2 \pmod{\ell-1}$, then

$$k(\rho) := \begin{cases} 2 & \text{if } \rho \text{ is finite flat,} \\ \ell+1 & \text{otherwise.} \end{cases}$$

2.4. Representations arising from elliptic curves

Theorem 2.10. *Suppose $A/\mathbf{Q}$ is a semistable elliptic curve and that $\rho_{A,\ell}$ is irreducible. Let Δ_A denote the minimal discriminant of A. The representation $\rho_{A,\ell}$ is finite flat at ℓ if and only if $\ell \mid \mathrm{ord}_\ell \Delta_A$. If $p \neq \ell$, then $\rho_{A,\ell}$ is unramified at p if and only if $\ell \mid \mathrm{ord}_p \Delta_A$.*

Proof. The first statement is Proposition 5 of [**102**].

When A has good reduction at p, the second statement holds (see Exercise 15). Suppose A has multiplicative reduction at p. There is an unramified extension K of $\mathbf{Q}_p$ such that A has split multiplicative reduction at p. Consider the Tate curve $\mathbf{G}_m/q^{\mathbf{Z}}$ over K associated to A. Thus $\overline{\mathbf{Q}}_p^*/q^{\mathbf{Z}} \cong A(\overline{\mathbf{Q}}_p)$ as $\mathrm{Gal}(\overline{\mathbf{Q}}_p/K)$-modules. The ℓ-torsion points $A[\ell]$ correspond to the points $\{\zeta_\ell^a (q^{1/\ell})^b : 0 \leq a, b < \ell\}$ in the Tate curve. The extension $K(\zeta_\ell, q^{1/\ell})$ of K is unramified because $\ell \neq p$ and $\mathrm{ord}_p(q) = \mathrm{ord}_p(\Delta_A)$ is divisible by ℓ. Since an unramified extension of an unramified extension is unramified, the extension $K(\zeta_\ell, q^{1/\ell})$ of $\mathbf{Q}_p$ is unramified, which proves the second part of the theorem. □

2.4.1. Frey curves

Using Theorem 2.10 we see that the Shimura-Taniyama conjecture together with Serre's conjecture implies Fermat's Last Theorem. Suppose (a, b, c) is a solution to the Fermat equation $a^\ell + b^\ell = c^\ell$ with $\ell \geq 11$ and $abc \neq 0$. Consider the Frey curve A given by the equation $y^2 = x(x - a^\ell)(x + b^\ell)$; it is an elliptic curve with discriminant $\Delta_A = \frac{((abc)^2)^\ell}{2^8}$. By [**93**, §4.1, Prop. 6] the representation $A[\ell]$ is irreducible. Theorem 2.10 implies that $\rho_{A,\ell}$ is unramified, except possibly at 2 and ℓ. Thus $N(\rho) \mid 2$, and $k(\rho) = 2$ since $\ell \mid \mathrm{ord}_\ell(\Delta_A)$. But there are no cusp forms of level 2 and weight 2. The modularity of A (proved in [**114, 117**]), together with the weak conjecture of Serre (enough of which is proved in [**84**]), leads to a contradiction.

2.4.2. Examples

Using Theorem 2.10 we can frequently determine the Serre invariants $N(\rho)$ and $k(\rho)$ of a representation ρ attached to an elliptic curve. When $N(\rho) < N$, it is illustrative to verify directly that there is a newform of level $N(\rho)$ that also gives rise to ρ. For example, there is a unique weight-2 normalized newform

$$f = q + q^2 - q^3 - q^4 - 2q^5 - q^6 + 4q^7 - 3q^8 + q^9 + \cdots$$

on $\Gamma_0(33)$. One of the elliptic curves associated to f is the curve A given by the equation

$$y^2 + xy = x^3 + x^2 - 11x.$$

The discriminant of A is $\Delta = 3^6 \cdot 11^2$ and the conductor is $N = 3 \cdot 11$. Because A is semistable and there are no elliptic curves 3-isogenous to A, the associated mod 3 representation $\rho = \rho_{A,3} : G_{\mathbf{Q}} \to \mathrm{Aut}(A[3])$ is surjective (see Section 1.4). Since $3 \mid \mathrm{ord}_3 \Delta_A$, the Serre weight and level are $k(\rho) = 2$ and $N(\rho) = 11$. As predicted by Serre's conjecture, there is a weight-2 newform on $\Gamma_0(11)$ such that if B is one of the three elliptic curves of conductor 11 (it does not matter which), then $B[3] \approx A[3]$ as representations of $G_{\mathbf{Q}}$. Placing the eigenforms corresponding to A and B next to each other, we observe that their Fourier coefficients are congruent modulo 3:

$$\begin{array}{rcccccccccccc} f_A &=& q & +q^2 & -q^3 & -q^4 & -2q^5 & -q^6 & +4q^7 & -3q^8 & +q^9 & + & \cdots \\ f_B &=& q & -2q^2 & -q^3 & +2q^4 & +q^5 & +2q^6 & -2q^7 & & -2q^9 & + & \cdots . \end{array}$$

Next consider the elliptic curve A cut out by the equation

$$y^2 + y = x^3 + x^2 - 12x + 2.$$

It has conductor $N = 141 = 3 \cdot 47$ and discriminant $\Delta = 3^7 \cdot 47$. Since $\mathrm{ord}_3(\Delta)$ is divisible by 7, the mod 7 representation $\rho_{A,7}$ has Serre invariants $k(\rho_{A,7}) = 2$ and $N(\rho_{A,7}) = 47$. In confirmation of Serre's conjecture, we find a form $f \in S_2(\Gamma_0(47))$ that gives rise to $\rho_{A,7}$. The Fourier coefficients of f generate a quartic field.

Next consider $\rho_{A,3}$, whose Serre invariants are $N(\rho_{A,3}) = 47$ and, since 3 does not divide $\mathrm{ord}_3(\Delta)$, $k(\rho_{A,3}) = \ell + 1 = 4$. In $S_4(\Gamma_0(47))$ there are two conjugacy classes of eigenforms, which are defined over fields of degree 3 and 8, respectively. The one that gives rise to $\rho_{A,3}$ is

$$g = q + aq^2 + (-1/2a^2 - 5/2a - 1)q^3 + (a^2 - 8)q^4 + (a^2 + a - 10)q^5 + \cdots,$$

where $a^3 + 5a^2 - 2a - 12 = 0$.

2.5. Companion forms

Suppose f is a newform of weight k with $2 \le k \le \ell + 1$. Let ℓ be an ordinary prime, so $a_\ell(f)$ is not congruent to 0 modulo a prime λ lying over ℓ and

$$\rho_{f,\lambda}|_{I_\ell} \sim \begin{pmatrix} \chi^{k-1} & * \\ 0 & 1 \end{pmatrix}.$$

Is this representation split or not? Put another way, can $*$ be taken equal to 0, after an appropriate choice of basis? For how many ℓ do these representations split? We suspect that the ordinary split primes ℓ are in the minority, among all primes. How can we quantify the number of split primes?

If $* = 0$, then

$$\rho|_{I_\ell} \sim \begin{pmatrix} 1 & 0 \\ 0 & \chi^{k-1} \end{pmatrix},$$

so

$$\rho|_{I_\ell} \otimes \chi^{\ell-k} \sim \begin{pmatrix} \chi^{\ell-k} & 0 \\ 0 & 1 \end{pmatrix}.$$

Assume that $2 \leq 1+\ell-k \leq \ell+1$, so $k(\rho \otimes \chi^{\ell-k}) = 1+\ell-k$. Using the θ-operator we see that $\rho \otimes \chi^{\ell-k}$ is modular, of *some* weight and level. To say that it is modular of Serre's conjectured weight $k(\rho)$ is to make a much strong statement. If $\rho \otimes \chi^{\ell-k}$ is indeed modular of weight $1+\ell-k$, then by definition there exists an eigenform g of weight $1+\ell-k$ with $\rho_g \sim \rho_f \otimes \chi^{\ell-k}$. Such an eigenform g, if it exists, is called a *companion* of f. The existence of g is far from obvious.

We can extend the notion of companion form to the case when $k(\rho) = \ell$. In this case the companion has weight 1. If ρ is unramified at ℓ, then we expect ρ to also arise from a weight-1 eigenform.

The existence of a companion form was proved (assuming unchecked compatibilities) in most cases in which $k < \ell$ by Gross in [**46**] and in a few cases when $k = \ell$. Using new methods, Coleman and Voloch [**17**] proved all cases except $k = \ell = 2$. The arguments of Coleman and Voloch do not require verification of Gross's unchecked compatibilities.

CHAPTER 3

Optimizing the level

Consider an irreducible Galois representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ that arises from a newform of weight k and level N. Serre defined integers $k(\rho)$ and $N(\rho)$, and conjectured that ρ arises from a newform of weight $k(\rho)$ and level $N(\rho)$. In Chapter 2 we sketched Edixhoven's proof that if $\ell \nmid N$ then ρ arises from an newform of weight $k(\rho)$ and level N. In this chapter, we introduce some of the techniques used in proving that ρ arises from a newform level $N(\rho)$. For more details, see [**84, 87**].

In [**102**, §1.2] Serre defined the *optimal level* $N(\rho)$ as the prime-to-ℓ part of the Artin conductor of ρ. Recall that $N(\rho)$ is a product $\prod p^{n(p)}$ over prime numbers $p \neq \ell$. The integer $n(p)$ is defined by restricting ρ to a decomposition group D_p at p. Consider the sequence of ramification groups $G_0 \supset G_1 \supset \cdots \supset G_i \supset \cdots$ where G_0 is the inertia subgroup I_p of D_p. Let V be a vector space over $\overline{\mathbf{F}}_\ell$ affording the representation ρ, and for each $i \geq 0$ let V_i be the subspace of V consisting of those $v \in V$ that are fixed by G_i. Then

$$n(p) := \sum_{i=0}^{\infty} \frac{1}{(G_0 : G_i)} \dim V/V_i.$$

3.1. Reduction to weight 2

The optimal level $N(\rho)$ is not divisible by ℓ. The first step in level optimization is to strip the power of ℓ from N. When ℓ is odd, this is done explicitly in [**87**, §2]; for the case $\ell = 2$ see [**9**, §1]. Many of the arguments and key ideas are due to Serre [**94**]. This proof that ℓ can be stripped from the level uses concrete techniques of Serre [**95**, §3], [**98**, Thm. 5.4], and Queen [**78**, §3]; it involves multiplying f by suitable Eisenstein series and taking traces. Katz's theory of ℓ-adic modular forms suggests an alternative method. A classical form of weight 2 and level $M\ell^m$ is an ℓ-adic form of level M; the mod ℓ reduction of this form is classical of level M and some weight, and is congruent to f. See the appendices of [**60**] and the discussions in [**49**, §1] and [**50**, §1].

The next step is to replace f by a newform of weight between 2 and $\ell + 1$ that gives rise to a twist of ρ. Twisting ρ by the mod ℓ cyclotomic character χ preserves N; this is because $\rho \otimes \chi$ arises from $\theta(f) = q\frac{d}{dq}(f)$, which also has level N. Theorem 2.7 asserts that some twist $\rho \otimes \chi^i$ of ρ arises from a form g of weight between 2 and $\ell + 1$. If $\rho \otimes \chi^i$ arises from a newform of level N, then ρ also arises from a newform of the same level, so we can replace f by g and k by the weight of g. By results discussed in Chapter 2, we may assume that $k = k(\rho \otimes \chi^i)$. For the case $\ell = 2$ see [**9**, Prop. 1.3(a)].

We have reduced to considering a representation ρ that arises from a newform f of weight $k(\rho)$ and level N not divisible by ℓ. The weight satisfies $2 \leq k(\rho) \leq \ell+1$, but N need not equal $N(\rho)$. That N is a multiple of $N(\rho)$ is a theorem proved by both Carayol [**12**] and Livné [**70**, Prop. 0.1].

In order to lower N it is convenient to work systematically with form of weight 2. Paradoxically, even though we have just taken all powers of ℓ out of N, we are now going to allow one power of ℓ back into N. This allows us to reduce to weight 2 and realize ρ as a group of torsion points on an abelian variety. An alternative approach (see [**41, 57**]) is to avoid this crutch and work directly with representations coming from arbitrary weights between 2 and $\ell+1$; these are realized in étale cohomology groups. This later approach has the advantage that $X_0(N)$ has good reduction at ℓ.

Reduction to weight 2 is accomplished using a general relationship that originates with ideas of Koike and Shimura. In characteristic ℓ, eigenforms of level N whose weights satisfy $2 < k \leq \ell+1$ correspond to eigenforms of weight 2 and level ℓN (see [**87**, Thm. 2.2]):

$$\left\{ \text{ } 2 < k \leq \ell+1, \text{ level } N \text{ } \right\} \leftsquigarrow \left\{ \text{ } k = 2, \text{ level } \ell N \text{ } \right\}.$$

Thus we can and do work with weight 2 and level

$$N^* := \begin{cases} N & \text{if } k = 2, \\ N\ell & \text{if } k > 2. \end{cases}$$

3.2. Geometric realization of Galois representations

To understand representations arising from modular forms, it is helpful to realize these representations inside of geometric objects such as $J := J_1(N^*)$. These representations are constructed geometrically with the help of the Hecke algebra

$$\mathbf{T} := \mathbf{Z}[\ldots T_n \ldots],$$

which was defined in Section 2.3. Recall that $\mathbf{T}$ is a commutative subring of $\mathrm{End}_{\mathbf{Q}} J$ that is free as a module over $\mathbf{Z}$, and that its rank is equal to the dimension of J. When N is cube free, $\mathbf{T}$ is an order in a product of integer rings of number fields; this is a result of Coleman and Edixhoven (see [**16**, Thm. 4.1]). In contrast, the Hecke operators T_p, for $p^3 \mid N$, are usually not semisimple (see Exercise 3).

It is fruitful to view a newform f as a homomorphism

$$\mathbf{T} \to \mathcal{O} = \mathbf{Z}[\ldots a_n \ldots], \qquad T_n \mapsto a_n.$$

Letting $\varphi : \mathcal{O} \to \overline{\mathbf{F}}_\ell$ be the map sending a_p to $\mathrm{tr}(\rho(\mathrm{Frob}_p)) \in \overline{\mathbf{F}}_\ell$, we obtain an exact sequence $0 \to \mathfrak{m} \to \mathbf{T} \to \overline{\mathbf{F}}_\ell$ with $\mathfrak{m}$ a maximal ideal.

Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be an irreducible Galois representation that arises from a weight-2 newform f. The next step, after having attached a maximal ideal $\mathfrak{m}$ to f and φ, is to find a $\mathbf{T}/\mathfrak{m}$-vector space affording ρ inside of the group of ℓ-torsion points of J. Following [**71**, §II.7], we consider the $\mathbf{T}/\mathfrak{m}$-vector space

$$J[\mathfrak{m}] := \{P \in J(\overline{\mathbf{Q}}) : tP = 0 \text{ all } t \in \mathfrak{m}\} \subset J(\overline{\mathbf{Q}})[\ell] \approx (\mathbf{Z}/\ell\mathbf{Z})^{2g}.$$

Since the endomorphisms in $\mathbf{T}$ are $\mathbf{Q}$-rational, $J[\mathfrak{m}]$ comes equipped with a linear action of $G_{\mathbf{Q}}$.

That $\mathrm{tr}(\rho(\mathrm{Frob}_p))$ and $\det(\rho(\mathrm{Frob}_p))$ both lie in the subfield $\mathbf{T}/\mathfrak{m}$ of $\overline{\mathbf{F}}_\ell$ suggests that ρ has a model over $\mathbf{T}/\mathfrak{m}$, in the sense that ρ is equivalent to a representation taking values in $\mathrm{GL}(2, \mathbf{T}/\mathfrak{m}) \subset \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$.

Lemma 3.1. *The representation ρ has a model $\rho_\mathfrak{m}$ over the finite field $\mathbf{T}/\mathfrak{m}$.*

Proof. This is a classical result of I. Schur. Brauer groups of finite fields are trivial (see e.g., [**100**, X.7, Ex. a]), so the argument of [**99**, §12.2] proves the lemma.

Alternatively, when the residue characteristic ℓ of $\mathbf{T}/\mathfrak{m}$ is odd, the following more direct proof can be used. Complex conjugation acts through ρ as a matrix with distinct $\mathbf{F}_\ell$-rational eigenvalues; another well known theorem of Schur [**90**, IX a] (cf. [**116**, Lemme I.1]) then implies that ρ can be conjugated into a representation with values in $\mathrm{GL}(2, \mathbf{T}/\mathfrak{m})$. □

3.3. Multiplicity one

Let $V_\mathfrak{m}$ be a vector space affording $\rho_\mathfrak{m}$. Under the assumption that $\rho_\mathfrak{m}$ is absolutely irreducible, Boston, Lenstra, and Ribet (see [**6**]) proved that $J[\mathfrak{m}]$ is isomorphic as a $G_\mathbf{Q}$-module to a sum of copies of $V_\mathfrak{m}$:

$$J[\mathfrak{m}] \approx \bigoplus_{i=1}^{t} V_\mathfrak{m}.$$

The number of copies of $V_\mathfrak{m}$ is called the *multiplicity* of $\mathfrak{m}$. When ℓ is odd, the hypothesis of irreducibility of $\rho_\mathfrak{m}$ is equivalent to absolute irreducibility (see Exercise 3).

Proposition 3.2. *The multiplicity t is at least* 1.

Proof. Let $\mathbf{T} \subset \mathrm{End}(J)$ be the Hecke algebra associated to J. Because $\mathbf{T} \otimes \mathbf{Z}_\ell$ is an algebra of finite rank over the local ring $\mathbf{Z}_\ell$, we have a decomposition

$$\mathbf{T} \otimes \mathbf{Z}_\ell = \bigoplus_{\lambda|\ell} \mathbf{T}_\lambda,$$

where λ runs through the maximal ideals of $\mathbf{T}$ lying over ℓ, and $\mathbf{T}_\lambda$ denotes the completion of $\mathbf{T}$ at λ (see, e.g., [**37**, Cor. 7.6]). The Tate module

$$\mathrm{Tate}_\ell J := \mathrm{Hom}(\mathbf{Q}_\ell/\mathbf{Z}_\ell, \cup_{n\geq 1} J[\ell^n]) \cong \varprojlim J[\ell^n]$$

is a free $\mathbf{Z}_\ell$-module of rank equal to twice the dimension of J. For each maximal ideal λ of $\mathbf{T}$ lying over ℓ, let $e_\lambda \in \mathbf{T} \otimes \mathbf{Z}_\ell$ denote the corresponding idempotent; thus $e_\lambda^2 = e_\lambda$ and $\sum_{\lambda|\ell} e_\lambda = 1$. The map $x \mapsto \sum_\lambda e_\lambda x$ gives a decomposition

$$\mathrm{Tate}_\ell J \xrightarrow{\cong} \bigoplus_{\lambda|\ell} e_\lambda \,\mathrm{Tate}_\ell J.$$

The ring $\mathrm{End}(J) \otimes \mathbf{Z}_\ell$ operates faithfully on $\mathrm{Tate}_\ell J$ (see, e.g., [**74**, Lem. 12.2]), so each summand $e_\lambda \,\mathrm{Tate}_\ell J$ is nonzero. Set

$$\mathrm{Tate}_\lambda J := \mathrm{Hom}(\mathbf{Q}_\ell/\mathbf{Z}_\ell, \cup_{n\geq 1} J[\lambda^n]).$$

We claim that $\mathrm{Tate}_\lambda J$ is identified with $e_\lambda \,\mathrm{Tate}_\ell J$ under the natural inclusion $\mathrm{Tate}_\lambda J \subset \mathrm{Tate}_\ell J$. Denote by $\tilde{\lambda}$ the maximal ideal in $\mathbf{T} \otimes \mathbf{Z}_\ell$ generated by λ. Let n be a positive integer, and let I be the ideal in $\mathbf{T}_\lambda$ generated by ℓ^n. Because $\mathbf{T}_\lambda$ is

a local ring with maximal ideal $\tilde{\lambda}$, there is an integer m such that $\tilde{\lambda}^m \subset I$. Since I is principal and generated by ℓ^n, and $\mathbf{T}$ acts on $e_\lambda J[\ell^n]$ through $\mathbf{T}_\lambda$, we have

$$e_\lambda J[\ell^n] = (e_\lambda J[\ell^n])[I] \subset (e_\lambda J[\ell^n])[\tilde{\lambda}^m] \subset (e_\lambda J[\ell^n])[\lambda^m] \subset J[\lambda^m].$$

This shows that $e_\lambda \operatorname{Tate}_\ell J \subset \operatorname{Tate}_\lambda J$. Next suppose $\lambda' \neq \lambda$ and let n be a positive integer. Since $\mathbf{T}_\lambda$ acts on $J[\lambda^n]$ through $\mathbf{T}/\lambda^n = \mathbf{T}_\lambda/\tilde{\lambda}^n$, we have $e_{\lambda'} J[\lambda^n] = 0$, so

$$J[\lambda^n] = \sum_{\text{all } \lambda'} e_{\lambda'} J[\lambda^n] = e_\lambda J[\lambda^n].$$

The other inclusion $\operatorname{Tate}_\lambda J = e_\lambda \operatorname{Tate}_\lambda J \subset e_\lambda \operatorname{Tate}_\ell J$, which we need to prove equality, then follows.

We apply the above conclusion with $\lambda = \mathfrak{m}$. Since $\operatorname{Tate}_\mathfrak{m} J \neq 0$, some $J[\mathfrak{m}^r]$ is nonzero; let r be the smallest such integer. Following [**71**, p. 112], observe that for each generating set of elements $a_1, \ldots, a_t$ of the $\mathbf{T}/\mathfrak{m}$-vector space $\mathfrak{m}^{r-1}/\mathfrak{m}^r$, the map $x \mapsto a_1 x \oplus \cdots \oplus a_t x$ is an injection of the module $J[\mathfrak{m}^r]/J[\mathfrak{m}^{r-1}]$ into the direct sum of t copies of $J[\mathfrak{m}]$. Thus $J[\mathfrak{m}]$ is nonzero. □

The special case $t = 1$, in which the multiplicity is one, plays a central role in the development of the theory. A detailed summary of multiplicity one results can be found in [**32**, §9], and some supplementary results are contained in [**117**, Thm. 2.1]. In general, the multiplicity can be greater than one (see [**72**, §13] and [**63**]).

3.3.1. Multiplicity one representations

Let $\rho : G_\mathbf{Q} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be an irreducible modular Galois representation such that

$$2 \leq k(\rho) \leq \ell + 1.$$

Consider pairs (N, α) where $N \geq 1$ is an integer with the property that $\ell \nmid N$ if $k(\rho) = 2$ and $\ell \mid\mid N$ if $k(\rho) > 2$, together with maps $\alpha : \mathbf{T}_N \to \overline{\mathbf{F}}_\ell$, such that $\alpha(T_p) = \operatorname{tr}(\rho(\mathrm{Frob}_p))$ and $\alpha(p\langle p\rangle) = \det(\rho(\mathrm{Frob}_p))$ for almost all p. Here $\mathbf{T}_N$ is the Hecke algebra associated to $S_2(\Gamma_1(N))$. Note that if (N, α) is such a pair and $\mathfrak{m} = \ker(\alpha)$, then

$$\rho \approx \rho_\mathfrak{m} \otimes_{\mathbf{T}/\mathfrak{m}} \overline{\mathbf{F}}_\ell,$$

where $\alpha : \mathbf{T}/\mathfrak{m} \hookrightarrow \overline{\mathbf{F}}_\ell$ and $\rho_\mathfrak{m}$ is the unique (up to isomorphism) semisimple representation over $\overline{\mathbf{F}}_\ell$ such that

$$\operatorname{tr}(\rho_\mathfrak{m}(\mathrm{Frob}_p)) = \alpha(T_p) \qquad \det(\rho_\mathfrak{m}(\mathrm{Frob}_p)) = \alpha(p\langle p\rangle)$$

for almost all p.

Definition 3.3. ρ is a multiplicity one representation if $J_1(N)[\ker \alpha]$ has dimension 2 for all pairs (N, α) as above.

Remark 3.4.
(1) If $J_1(N)[\ker \alpha]$ has dimension 2 then $\rho_\mathfrak{m} = J_1(N)[\ker \alpha]$ by Eichler-Shimura, see [**6**].
(2) The definition extends to arbitrary modular Galois representations ρ as follows. As explained in Section 2.2, every ρ has a twist $\rho \otimes \chi^i$ by some power of the cyclotomic character such that $k(\rho \otimes \chi^i) \leq \ell + 1$. We say that ρ is a *multiplicity one representation* if $\rho \otimes \chi^i$ is a multiplicity one representation.

3.3.2. Multiplicity one theorems

Techniques for proving multiplicity one results were pioneered by Mazur in [**71**] who considered $J_0(p)$ with p prime. Let f be an eigenform and fix a nonzero prime λ of the ring generated by the Fourier coefficients of f such that $\rho_{f,\lambda}$ is absolutely irreducible. View the Hecke algebra $\mathbf{T}$ as a subring of $\mathrm{End}(J_0(p))$, and let $\mathfrak{m}$ be the maximal ideal associated to f and λ. Let $V_{\mathfrak{m}}$ again be a two-dimensional $\mathbf{T}/\mathfrak{m}$-vector space that affords $\rho_{\mathfrak{m}} : G_{\mathbf{Q}} \to \mathrm{GL}(2, \mathbf{T}/\mathfrak{m})$. Mazur proved (see Prop. 14.2, ibid.) that $J[\mathfrak{m}] \approx V_{\mathfrak{m}}$, except perhaps when $\mathfrak{m}$ is ordinary of residue characteristic $\ell = 2$. The missing ordinary case can be treated under suitable hypothesis. If $\rho_{\mathfrak{m}}$ restricted to a decomposition group at 2 is not contained in the scalar matrices, then $J[\mathfrak{m}] \approx V_{\mathfrak{m}}$ (see, e.g., [**9**, Prop. 2.4]). The results of Mazur are extended in [**72**] and [**84**, §5].

Theorem 3.5. *An irreducible modular Galois representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ is a multiplicity one representation, except perhaps when all of the following hypothesis on ρ are simultaneously satisfied:*

- *$k(\rho) = \ell$;*
- *ρ is unramified at ℓ;*
- *ρ is ordinary at ℓ;*
- *$\rho|_{D_\ell} \sim \left(\begin{smallmatrix} \alpha & * \\ 0 & \beta \end{smallmatrix}\right)$ with $\alpha = \beta$.*

Proof. See [**32**, §9], [**117**, Thm. 2.1], and [**9**, Prop. 2.4] for the case $\ell = 2$. □

In [**46**, §12] Gross proves multiplicity one when $\alpha \neq \beta$, $k(\rho) \leq \ell$, and ρ is ordinary; he uses this result in his proof of the existence of companion forms. In contrast, Coleman and Voloch [**17**] prove the existence of companion forms when $\alpha = \beta$ and $\ell > 2$ using a method that avoids the need for multiplicity one.

Remark 3.6. L. Kilford of London, England has recently discovered an example at prime level 503 in which multiplicity one fails. Let E_1, E_2, and E_3 be the three elliptic curves of conductor 503, and for each $i = 1, 2, 3$, let $\mathfrak{m}_i$ be the maximal ideal of $\mathbf{T} \subset \mathrm{End}(J_0(503))$ generated by 2 and all $T_p - a_p(E_i)$, with p prime. Each of the Galois representations $E_i[2]$ is irreducible, and one can check that $\mathfrak{m}_1 = \mathfrak{m}_2 = \mathfrak{m}_3$. If multiplicity one holds, then $E_1[2] = E_2[2] = E_3[2]$ inside of $J_0(503)$. However, this is not the case, as a modular symbols computation in the integral homology $H_1(X_0(N), \mathbf{Z})$ reveals that $E_1 \cap E_2 = \{0\}$.

3.3.3. Multiplicity one for mod 2 representations

For future reference, we now wish to consider multiplicity one in the following rather extreme situation. Suppose that $\ell = 2$, and let ρ be a mod ℓ representation arising from a form of weight either 2 or 3. If the weight is 3 then ρ is not finite at 2; this can be used to deduce multiplicity one by adapting the arguments of [**72**] (see the proof of [**9**, Prop. 2.4]). When the weight is 2, we have the following proposition.

Proposition 3.7. *Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_2)$ be an irreducible Galois representation that arises from a weight-2 form $f = \sum a_n q^n$ on $\Gamma = \Gamma_1(N) \cap \Gamma_0(2)$ with N odd, and let ε be the character of f. If $\overline{a}_2^2 \not\equiv \overline{\varepsilon}(2) \in \overline{\mathbf{F}}_2$, then ρ is a multiplicity one representation.*

Proof. Let $\mathfrak{m}$ be the maximal ideal associated to f in the Hecke algebra $\mathbf{T}$ attached to Γ. Because the weight of f is 2, the representation ρ is finite at 2. If ρ is supersingular then the inertia group I_2 operates through the two fundamental characters of level 2. These both have order $\ell^2 - 1 = 3 \neq 1$, so ρ is ramified and this can be used to deduce multiplicity one. If ρ is ordinary then $\rho|_{D_2} \sim \left(\begin{smallmatrix} \alpha & * \\ 0 & \beta \end{smallmatrix}\right)$ with β unramified and $\beta(\mathrm{Frob}_2) \equiv T_2 \bmod \mathfrak{m}$. The determinant $\alpha\beta$ of $\rho|_{D_2}$ is $\chi \cdot \varepsilon$ where χ is the mod 2 cyclotomic character and ε is unramified at 2. Since χ, ε, and β are unramified, α is also unramified. Since $\chi(\mathrm{Frob}_2) = 1$ and $\alpha\beta = \chi\varepsilon$, we have $\alpha(\mathrm{Frob}_2) = \beta^{-1}(\mathrm{Frob}_2)\varepsilon(2) = a_2^{-1}\varepsilon(2) \pmod{\mathfrak{m}}$. The further condition, under which we might not know multiplicity one, is $\alpha|_{D_2} = \beta|_{D_2}$; expressed in terms of the image of Frobenius, this becomes $a_2^{-1}\varepsilon(2) \equiv a_2 \pmod{\mathfrak{m}}$, or equivalently, $a_2^2 \equiv \varepsilon(2) \pmod{\mathfrak{m}}$. By hypothesis, this latter condition does not hold. □

3.4. The key case

We have set our problem up so that level optimization pertains to weight-2 forms of appropriate level, and takes place on Jacobians of modular curves. This level optimization problem was described, and partially treated, in a paper of Carayol [**12**]. In this paper, Carayol reduced the problem to the following key case.

Key case: Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be a Galois representation that arises from a weight-2 newform f of level pM, with $p \nmid \ell M$, and character $\varepsilon : (\mathbf{Z}/pM\mathbf{Z})^* \to \mathbf{C}^*$. Assume that ρ is unramified at p, and that ε factors through the natural map $(\mathbf{Z}/pM\mathbf{Z})^* \to (\mathbf{Z}/M\mathbf{Z})^*$. Show that ρ arises from a form of level M.

In the key case, the character ε of f is unramified at p. Thus f, a priori on $\Gamma_1(pM)$, is also on the bigger group $\Gamma_1(M) \cap \Gamma_0(p)$; that is, f lies in $S_2(\Gamma_1(M) \cap \Gamma_0(p))$.

Example 3.8. Consider the representation ρ arising from the 7-division points of the modular elliptic curve A of conductor $N_A = 3 \cdot 47$ and minimal discriminant $\Delta_A = 3^7 \cdot 47$. (The curve A is labeled **141A** in Cremona's notation [**20**].) The newform f corresponding to A is on $\Gamma_0(3{\cdot}47)$. As in Section 1.4, since $\mathrm{ord}_3(\Delta_A) = 7$, the representation ρ is unramified at 3 and $N(\rho) = 47$. To optimize the level means to find a form g on $\Gamma_0(47)$ that gives rise to ρ.

Example 3.9 (Frey curves)**.** The elliptic curves that Frey associated in [**42**] to hypothetical solutions of the Fermat equation $x^\ell + y^\ell = z^\ell$ give rise to mod ℓ Galois representations. According to Wiles's theorem [**117**], there is a weight-2 form f of level $2L$, with L big and square free, that gives rise to ρ. At the same time, $N(\rho) = 2$. Taking p to be any odd prime dividing L, we are put in the key case. If we can optimize the level, then we eventually reach a contradiction and thus deduce Fermat's Last Theorem.

The key case divides into two subcases; the more difficult one occurs when the following conditions are both satisfied:

— $p \equiv 1 \pmod{\ell}$;
— $\rho(\mathrm{Frob}_p)$ is a scalar matrix.

The second condition makes sense because $p \nmid N(\rho)$; since $\det(\rho(\mathrm{Frob}_p)) = \chi^{k-1}\varepsilon$, we know the scalar up to ± 1. The complementary case is easier; it can be treated

using "Mazur's principle" (see Section 3.9). Though Example 3.8 falls into the easier case because $3 \not\equiv 1 \pmod 7$, the proof of Fermat's Last Theorem requires level optimization in both cases.

Consider a modular representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ that arises from a newform of level N and weight $k = k(\rho)$, and assume that $\ell \nmid N$. The goal of level optimization is to show that there is a newform of Serre's optimal level $N(\rho)$ that gives rise to ρ.

As discussed in Section 3.1, ρ arises from a newform $f = \sum a_n q^n$ on $\Gamma_1(N^*)$ of weight 2 and some character ε. Thus there is a homomorphism φ from $\mathcal{O} = \mathbf{Z}[\ldots a_n \ldots]$ to $\overline{\mathbf{F}}_\ell$ such that $\varphi(a_p) = \mathrm{tr}(\rho(\mathrm{Frob}_p))$ for all $p \nmid \ell N^*$. Let $\mathbf{T}$ be the Hecke algebra associated to $S_2(\Gamma_1(N^*))$. The maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ associated to ρ is the kernel of the map sending T_n to $\varphi(a_n)$. As was discussed in the previous chapter, the representation ρ is realized geometrically inside the subspace $J[\mathfrak{m}] \subset J[\ell]$ of the ℓ-torsion of the Jacobian J of $X_1(N^*)$.

Problem. Fix a divisor p of $N^*/N(\rho)$. Find a newform whose level is a divisor of N^*/p that also gives rise to ρ.

Lemma 3.10. *Let ρ be as above, and suppose p is a prime such that $p \mid N^*$ but $p \nmid \ell N(\rho)$, so ρ is unramified at p. Let ε_p denote the p part of ε. Then either $\varepsilon_p = 1$ or $p \equiv 1 \pmod \ell$.*

Proof. The character ε is initially defined as a homomorphism $(\mathbf{Z}/N^*\mathbf{Z})^* \to \mathcal{O}^*$; the reduction $\overline{\varepsilon}$ is obtained by composing ε with $\varphi : \mathcal{O} \to \overline{\mathbf{F}}_\ell$. Since ρ is unramified at p, the determinant $\det(\rho) = \chi_\ell^{k-1}\overline{\varepsilon} = \chi_\ell \overline{\varepsilon}$ is also unramified at p. Because χ_ℓ is ramified only at ℓ, the character $\overline{\varepsilon}$ is unramified at p. Let $M = N^*/p^r$ where $r = \mathrm{ord}_p(N^*)$, and write $(\mathbf{Z}/N^*\mathbf{Z})^* \cong (\mathbf{Z}/p^r\mathbf{Z})^* \times (\mathbf{Z}/M\mathbf{Z})^*$. By restricting ε to each factor, we write ε as a product of two characters: $\varepsilon = \varepsilon_p \cdot \varepsilon^{(p)}$ where ε_p is a character of $(\mathbf{Z}/p^r\mathbf{Z})^*$ and $\varepsilon^{(p)}$ is a character of $(\mathbf{Z}/M\mathbf{Z})^*$. The character $\varepsilon^{(p)}$ has conductor dividing M, so it is unramified at p. By class field theory, ε_p is totally ramified at p, so the reduction $\overline{\varepsilon}$ is unramified at p precisely when $\overline{\varepsilon}_p = 1$; equivalently, $\overline{\varepsilon}$ is unramified at p exactly when ε_p has order a power of ℓ. If ε_p is non-trivial, then, since the order of ε_p divides the order $p^{r-1}(p-1)$ of a generator of $(\mathbf{Z}/p^r\mathbf{Z})^*$, a power of ℓ divides $p^{r-1}(p-1)$, so $p \equiv 1 \pmod \ell$ since $\ell \neq p$. □

In addition to his conjectures about the optimal weight and level, Serre also made a conjecture about the optimal character of a form giving rise to ρ. Let p be a prime not dividing $\ell N(\rho)$. Serre's optimal character conjecture implies that ρ, which we know to arise from a form on $\Gamma_1(M) \cap \Gamma_1(p^r)$, arises from a form on $\Gamma_1(M) \cap \Gamma_0(p^r)$, and this has been proved in most cases.

3.5. Approaches to level optimization in the key case

As discussed in Section 3.4, results of Carayol and Livné (see [**12, 70**]) reduce the level optimization problem to the following key case. The weight-2 newform f, a priori on $\Gamma_1(N^*)$, is in fact on the bigger group $\Gamma_1(M) \cap \Gamma_0(p)$, where $Mp = N^*$, $p \nmid M$, and ρ is unramified at p. The goal is to show that ρ arises from a newform on $\Gamma_1(M)$. This has been achieved when ℓ is odd, and in many cases when $\ell = 2$, using several level optimization techniques.

I. **Mazur's principle**
If either $\rho(\mathrm{Frob}_p)$ is not a scalar matrix or $p \not\equiv 1 \pmod{\ell}$, then an argument of Mazur, explained in Section 3.9, can be used to optimize the level.

II. **Multiplicity one**
It is possible to optimize the level if ρ is a multiplicity one representation, as explained in [**84, 9**] and Section 3.11. The cases in which multiplicity one is known were reviewed in Section 3.3. In particular, we do not know multiplicity one in some cases when $k(\rho) = \ell$ and the eigenvalues of Frob_p are not distinct.

III. **Using a pivot**
Suppose that M can be written as a product $M = qK$ with q a prime not dividing pK, that ρ arises from a form on $\Gamma_1(K) \cap \Gamma_0(pq)$, and that ρ is ramified at q and unramified at p. Then q can be used as a "pivot" to remove p from the level. This approach grew out of [**83**], and was introduced in the short paper [**86**]. In Section 3.10 we describe the approach and discuss the terminology.

IV. **Without multiplicity one**
When ℓ is odd and $\varepsilon = 1$, the level optimization theorem was proved in [**87**] using an argument that does not require ρ to have multiplicity one. The hypothesis $\ell \neq 2$ is used in the proof of Proposition 7.8 of [**87**] to force splitting of a short exact sequence. In [**26**], Diamond extended the results of [**87**] to cover the case of arbitrary character, still under the assumption that ℓ is odd. One encounters seemingly insurmountable difficulties in trying to push this argument through when $\ell = 2$.

3.6. Some commutative algebra

In this section we set up some of the commutative algebra that is required in order to lower levels. There are two injective maps

$$S_2(\Gamma_1(M)) \begin{array}{c} \hookrightarrow \\ \hookrightarrow \end{array} S_2(\Gamma_1(M) \cap \Gamma_0(p)) \ .$$

One is the inclusion $f(q) \mapsto f(q)$ and the other is $f(q) \mapsto f(q^p)$ (see Exercise 18). The *p-new subspace* $S_2(\Gamma_1(M) \cap \Gamma_0(p))^{p\text{-new}}$ is the complement, with respect to the Petersson inner product, of the subspace $\mathcal{S}$ generated by the two images of $S_2(\Gamma_1(M))$. The p-new subspace can also be defined algebraically as the kernel of the natural map from $S_2(\Gamma_1(M) \cap \Gamma_0(p))$ to the direct sum of two copies of $S_2(\Gamma_1(M))$.

Let $\mathbf{T}$ denote the Hecke algebra acting on $S_2(\Gamma_1(M) \cap \Gamma_0(p))$. If $p \nmid M$, then T_p acts on $\mathcal{S}$ as a direct sum of two copies of its action on $S_2(\Gamma_1(M))$; otherwise, T_p usually does not act diagonally (see Exercise 19). The image of $\mathbf{T}$ in $\mathrm{End}(\mathcal{S})$ is a quotient $\overline{\mathbf{T}}$ called the *p-new quotient*. A representation ρ associated to a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ arises from level M if and only if $\mathfrak{m}$ arises by pullback from a maximal ideal of $\overline{\mathbf{T}}$. Because the map $\mathbf{T} \to \overline{\mathbf{T}}$ is surjective, $\mathfrak{m}$ arises from level M if and only if the image of $\mathfrak{m}$ in$\overline{\mathbf{T}}$ is not the unit ideal (see Exercise 21).

3.7. Aside: Examples in characteristic two

Sections 3.7 and 3.8 can be safely skipped on a first reading.

To orient the reader, we focus for the moment on mod 2 representations that arise from elliptic curves. We give examples in which one of the level optimizations methods applies but the others do not. We do not consider method **IV** because it is not applicable to mod 2 representations. The hypothesis of the "multiplicity one" method **II** when $\ell = 2$ are discussed after the statement of Theorem 3.19 in Section 3.11. We were unable to find an example in which none of the level optimization theorems applies.

We will repeatedly refer to the following theorem, which first appeared in [**85**].

Theorem 3.11. *Suppose ρ arises from a newform in $S_2(\Gamma_0(N))$. Let $p \nmid \ell N$ be a prime satisfying one or both of the identities*

$$\operatorname{tr} \rho(\mathrm{Frob}_p) = \pm(p+1) \pmod{\ell}.$$

Then ρ arises from a newform of level pN.

3.7.1. III applies but I and II do not

In this section we give a mod 2 representations in which the pivot hypothesis of **III** is satisfied, but the hypotheses of **I** and **II** are not. Our example is obtained by applying Theorem 3.11 to the mod 2 representation attached to a well-chosen elliptic curve.

We will find an elliptic curve E of conductor $M = qR$ such that $\rho = E[2]$ is absolutely irreducible, ramified at q, unramified at 2, and $\rho(\mathrm{Frob}_2) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Because of the last condition, [**9**, Prop. 2.4] does not imply that ρ is a multiplicity one representation, so **II** does not apply. (In fact, following Remark 3.6, one sees that ρ is not a multiplicity one representation.) Likewise, **I** does not apply because $\rho(\mathrm{Frob}_2)$ is a scalar and the p we will chose will satisfy $p \equiv 1 \pmod 2$. Next we choose a prime $p \nmid 2qR$ such that $\rho_{E,2}(\mathrm{Frob}_p) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Let f be the newform associated to E. By Theorem 3.11 there is a newform g of level pqR such that

$$\rho_{g,\lambda} \approx \rho_{E,2}.$$

In particular,

$$\rho_{g,\lambda}(\mathrm{Frob}_p) = \rho_{E,2}(\mathrm{Frob}_p) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

is scalar and $p \equiv 1 \pmod 2$, so **I** does not apply. However, method **III** does apply with q used as a pivot.

For example, consider the elliptic curve E defined by the equation

$$y^2 + xy = x^3 - x^2 + 19x - 32.$$

The conductor of E is $N = 19 \cdot 109$, and the discriminant of the field $K = \mathbf{Q}(E[2])$ is $-19^3 \cdot 109^3$. We select $q = 19$ as our pivot. The prime $p = 73$ splits completely in K, so

$$\rho_{E,2}(\mathrm{Frob}_p) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

By Theorem 3.11 there is a form g of level $109 \cdot 19 \cdot 73$ that is congruent to the newform f attached to E modulo a prime lying over 2. Method **III** can be used to optimize the level, but neither method **I** nor **II** applies.

3.7.2. II applies but I and III do not

We exhibit a mod 2 representation for which method **II** can be used to optimize

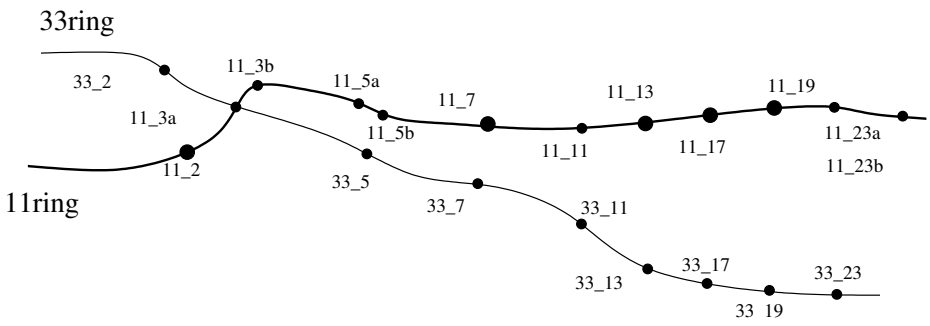


Figure 1. The spectrum of $\mathbf{T} \subset \mathrm{End}(S_2(\Gamma_0(33)))$, with $x = T_3$

the level, but neither method **I** nor **III** applies. Let K be the $\mathrm{GL}_2(\mathbf{F}_2)$-extension of $\mathbf{Q}$ obtained by adjoining all cube roots of 2. Then $K = \mathbf{Q}(E[2])$, where E is the elliptic curve $X_0(27)$ given by the equation $y^2 + y = x^3 - 7$. The prime $p = 31$ splits completely in K, so by Theorem 3.11 there is a newform f of level $31 \cdot 27$ and a maximal ideal λ of the appropriate Hecke algebra such that $\rho_{f,\lambda} \approx E[2]$. Neither method **I** nor **III** can be used to optimize the level of $\rho_{f,\lambda}$. Method **I** doesn't apply because 31 is odd and $\rho_{f,\lambda}(\mathrm{Frob}_{31}) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$; method **III** doesn't apply because the only odd prime that is ramified in K is 3, which does not exactly divide $31 \cdot 27$. If D_2 is a decomposition group at 2 then D_2 has image in $\mathrm{GL}_2(\mathbf{F}_2)$ of order 2, so it is not contained in the scalar matrices and **II** can be used to optimize the level of $\rho_{f,\lambda}$.

3.8. Aside: Sketching the spectrum of the Hecke algebra

It is helpful to understand the Hecke algebra geometrically using the language of schemes (see, e.g., [**38**]). The topological space underlying the scheme $\mathrm{Spec}(\mathbf{T})$ is the set of prime ideals of $\mathbf{T}$ endowed with the Zariski topology, in which the closed sets are the set of prime ideals containing a fixed ideal.

We can draw $\mathrm{Spec}(\mathbf{T})$ by sketching a diagram whose irreducible components correspond to the Galois conjugacy classes of eigenforms, and whose intersections correspond to congruences between eigenforms. When the level is not cube free, $\mathbf{T}$ can contain nilpotent elements, and then one might wish to include additional information. If $\sum a_n q^n$ is an eigenform, then the failure of $\mathbf{Z}[\ldots a_n \ldots]$ to be integrally closed can be illustrated by drawing singular points on the corresponding irreducible component; however, we do not do this below.

Example 3.12. The spectrum of the Hecke algebra associated to $\Gamma_0(33)$ is illustrated in Figure 1. The Hecke algebra $\mathbf{T} \subset S_2(\Gamma_0(33))$ has discriminant -99, as does the characteristic polynomial of T_3, so

$$\mathbf{T} = \mathbf{Z}[T_3]/((T_3 + 1)(T_3^2 + T_3 + 3)) \cong \mathbf{Z}[x]/((x + 1)(x^2 + x + 3)).$$

We sketch a curve corresponding to each of the two irreducible components. Some of the closed points (maximal) ideals are represented as dots. One component corresponds to the unique newform on $\Gamma_0(33)$, and the other corresponds to the two images of the newform on $\Gamma_0(11)$.

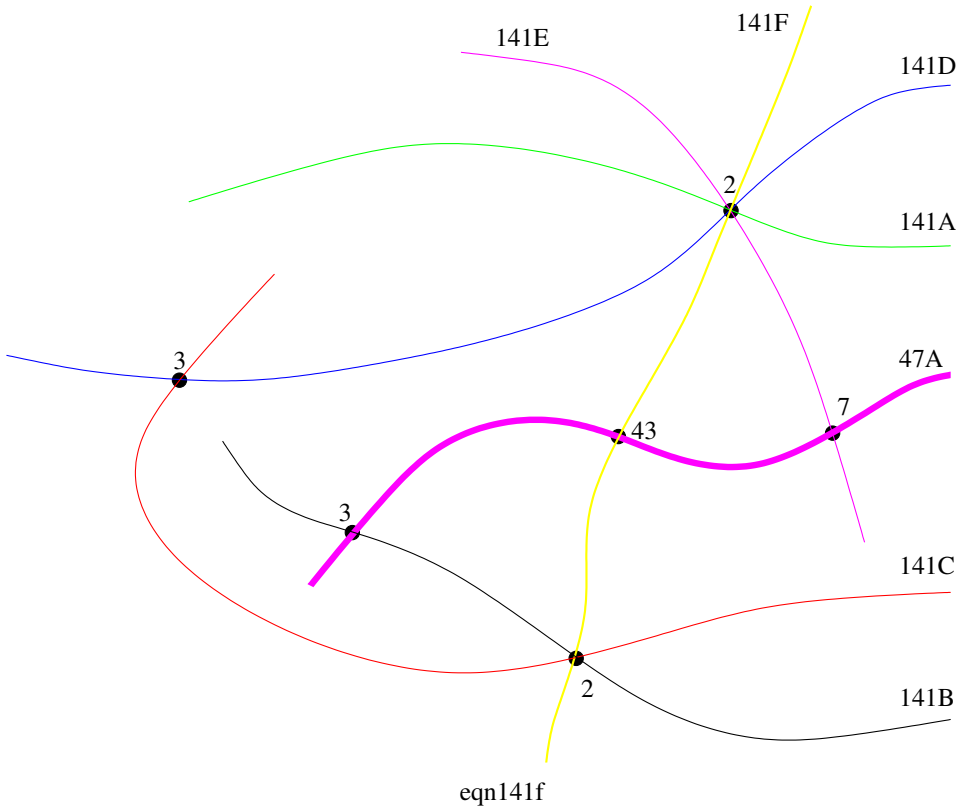


Figure 2. The spectrum of $\mathbf{T} \subset \mathrm{End}(S_2(\Gamma_0(141)))$

Example 3.13. Figure 2 is a diagram of the Hecke algebra associated to $S_2(\Gamma_0(3 \cdot 47))$. We have labeled fewer closed points than in Figure 1. The components are labeled by their isogeny class and the level at which they are new (the notation extends that of [**20**]). The component labeled **141F** corresponds to an eigenform whose Fourier coefficients generate a quadratic extension of $\mathbf{Q}$.

The newform corresponding to the elliptic curve A from Example 3.8 is labeled **141A**. Geometrically, the assertion that the level of $\rho_{A,7}$ can be optimized is represented by the characteristic-7 intersection between the component labeled **141A** and the old component **47A** coming from the unique Galois conjugacy class of newforms on $\Gamma_0(47)$.

3.9. Mazur's principle

A principle due to Mazur can be used to optimize the level in the key case, provided that a mild hypothesis is satisfied. The principle applies whenever $p \not\equiv 1 \pmod{\ell}$ and also in the case when $p \equiv 1 \pmod{\ell}$ but $\rho(\mathrm{Frob}_p)$ is not a scalar. This principle first appeared in [**84**, §6], then in [**26**, §4], and most recently when $\ell = 2$ in [**9**, pg. 7].

Theorem 3.14 (Mazur's Principle)**.** *Suppose that $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ arises from a newform f of weight 2 and level Mp, with $p \nmid M$, and character ε of conductor dividing M. Assume that ρ is unramified at p and that either $\rho(\mathrm{Frob}_p)$ is not a scalar matrix or $p \not\equiv 1 \pmod{\ell}$. Then ρ arises from a modular of level dividing M.*

We will require the following basic fact later in the proof.

Lemma 3.15 (Li)**.** *Let $f = \sum a_n q^n$ be a newform on $\Gamma_1(M) \cap \Gamma_0(p)$ of weight k. Then $a_p^2 = \varepsilon(p) p^{k-2}$.*

Proof. Li's proof is an easy application of her generalization to Γ_1 of the Atkin-Lehner theory of newforms [**69**, Thm. 3(iii)]. The newform f is an eigenvector for the operator W_p which is defined on $S_k(\Gamma_1(M) \cap \Gamma_0(p))$ by

$$W_p(f) = p^{k/2} f\left(\frac{apz + b}{Mpz + p}\right),$$

where a and b are integers such that $ap^2 - bMp = p$. By [**69**, Lem. 3],

$$g := T_p(f) + p^{k/2-1} W_p(f)$$

lies in $S_k(\Gamma_1(M))$. For all primes $q \nmid Mp$, the eigenvalue of T_q on the oldform g is the same as the eigenvalue of T_q on the newform f, so $g = 0$. By [**69**, Lem. 2] ${W_p}^2(f) = \varepsilon(p) f$, so $a_p^2 = \varepsilon(p) p^{k-2}$. □

Remark 3.16. The case of Lemma 3.15 that we will need can also be understood in terms of the local representation $\rho|_{G_p}$, which resembles the mod ℓ representation attached to a Tate curve, in the sense that $\rho|_{G_p} \sim \left(\begin{smallmatrix} \alpha\chi & * \\ 0 & \alpha \end{smallmatrix}\right)$. Our hypothesis include the assumption that ρ is unramified at p, so the two characters $\alpha\chi$ and α are unramified at p. Thus $\alpha(\mathrm{Frob}_p)$ makes sense; we have $\alpha(\mathrm{Frob}_p) = \overline{a}_p(f)$ and $\alpha\chi(\mathrm{Frob}_p) = \overline{a}_p(f)p$. Since $\det(\rho|_{G_p}) = \alpha^2\chi = \overline{\varepsilon}\chi$, we see that

$$\overline{a}_p^2 = \overline{\varepsilon}(p).$$

This local analysis of ρ was vastly generalized by Langlands in [**67**], which extends the analysis to include many ℓ-adic representations of possibly higher weight. See also [**13**].

Let $\mathbf{T}$ be the Hecke algebra associated to $\Gamma_1(M) \cap \Gamma_0(p)$, and let $\mathfrak{m}$ be the kernel of the following map $\mathbf{T} \to \overline{\mathbf{F}}_\ell$:

$$0 \longrightarrow \mathfrak{m} \longrightarrow \mathbf{T} \xrightarrow{T_n \mapsto \overline{a}_n,\ \langle d \rangle \mapsto \overline{\varepsilon}(d)} \overline{\mathbf{F}}_\ell.$$

As in Lemma 3.1, the determinants and traces of elements in the image of $\rho = \rho_{\mathfrak{m}}$ lie in $\mathbf{T}/\mathfrak{m} \subset \overline{\mathbf{F}}_\ell$, so there is a vector space $V \approx (\mathbf{T}/\mathfrak{m})^{\oplus 2}$ that affords $\rho_{\mathfrak{m}}$.

Next we realize $\rho_{\mathfrak{m}}$ as a group of division points in a Jacobian. The curve $X_1(Mp)$ corresponding to $\Gamma_1(Mp)$ covers the curve $X_1(M, p)$ corresponding to $\Gamma_1(M) \cap \Gamma_0(p)$. The induced map $J = \mathrm{Jac}(X_1(M, p)) \to J_1(Mp) = \mathrm{Jac}(X_1(Mp))$ has a finite kernel on which the Galois action is abelian.

Just as in Section 2.3.1.1, the Hecke algebra associated to $\Gamma_1(M) \cap \Gamma_0(p)$, can be constructed as a ring of correspondences on $X_1(M, p)$, then viewed as a subring $\mathbf{T} \subset \mathrm{End}_{\mathbf{Q}}(J)$. Inside of J we find the nonzero $G_{\mathbf{Q}}$-module $J[\mathfrak{m}] \approx \oplus_{i=1}^t V$. For the purposes of this discussion, we do not need to know that $J[\mathfrak{m}]$ is a direct sum of copies of V. The following weaker assertion, known long ago to Mazur [**71**, §14, pg. 112], will suffice: *$J[\mathfrak{m}]$ is a successive extension of copies of V.* In particular, $V \subset J[\mathfrak{m}]$. A weaker conclusion, true since $\ell \in \mathfrak{m}$, is that $V \subset J[\ell]$,

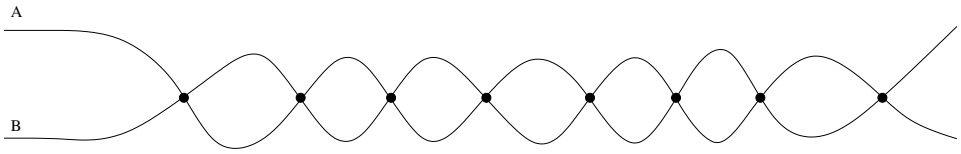


Figure 3. The reduction mod p of the Deligne-Rapoport model of $X_1(M,p)$

Our hypothesis that ρ is unramified at p translates into the inclusion $V \subset J[\ell]^{I_p}$, where I_p is an inertia group at p. By [**104**, Lem. 2], if A is an abelian variety over $\mathbf{Q}$ with good reduction at p, then $A[\ell]^{I_p} \cong A_{\mathbf{F}_p}[\ell]$. However, the modular curve $X_1(M,p)$ has bad reduction at p, so J is likely to have bad reduction at p—in this case it does. We are led to consider the Néron model $\mathcal{J}$ of J (see, e.g., [**5**]), which is a smooth commutative group scheme over $\mathbf{Z}$ satisfying the following property: the restriction map $\mathrm{Hom}_{\mathbf{Z}}(\mathcal{S},\mathcal{J}) \longrightarrow \mathrm{Hom}_{\mathbf{Q}}(\mathcal{S}_{\mathbf{Q}}, J)$ is bijective for all smooth schemes $\mathcal{S}$ over $\mathbf{Z}$. Passing to the scheme-theoretic closure, we have, inside of $\mathcal{J}$, a two-dimensional $\mathbf{T}/\mathfrak{m}$-vector space scheme $\mathcal{V}$.

In Section 2.3.1.1 we only defined $X_1(M,p)$ as a scheme over $\mathbf{Z}[1/Mp]$. Deligne and Rapoport [**25**] extended $X_1(M,p)$ to a scheme over $\mathbf{Z}[1/M]$ and computed the reduction modulo p. The introduction to [**62**] contains a beautiful historical discussion of the difficulties involved in extending modular curves over $\mathbf{Z}$.

We know a great deal about the reduction of $X_1(M,p)$ at p, which is frequently illustrated by the squiggly diagram in Figure 3. This reduction is the union of 2 copies of $X_1(M)_{\mathbf{F}_p}$ intersecting transversely at the supersingular points.

The subspace $S_2(\Gamma_1(M)) \oplus S_2(\Gamma_1(M))$ of $S_2(\Gamma_1(M) \cap \Gamma_0(p))$ is stable under the Hecke algebra $\mathbf{T}$, so there is a map $\mathbf{T} \to \mathrm{End}(S_2(\Gamma_1(M)) \oplus S_2(\Gamma_1(M)))$. The *$p$-old quotient* of $\mathbf{T}$ is the image $\overline{\mathbf{T}}$. Since the map $\mathbf{T} \to \overline{\mathbf{T}}$ is surjective, the image of $\mathfrak{m}$ in $\overline{\mathbf{T}}$ is an ideal $\overline{\mathfrak{m}}$. To optimize the level in the key case amounts to showing that $\overline{\mathfrak{m}}$ is not the unit ideal.

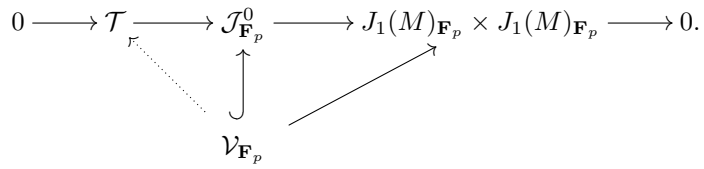
As is well known (cf. [**71**, Appendix, Prop 1.4]), the results of M. Raynaud [**82**] and Deligne-Rapoport [**25**] combine to produce an exact sequence

$$0 \longrightarrow \mathcal{T} \longrightarrow \mathcal{J}^0_{\mathbf{F}_p} \longrightarrow J_1(M)_{\mathbf{F}_p} \times J_1(M)_{\mathbf{F}_p} \longrightarrow 0, \tag{3.1}$$

where $\mathcal{T}$ is a torus, i.e., $\mathcal{T}_{\overline{\mathbf{F}}_p} \approx \mathbf{G}_m \times \cdots \times \mathbf{G}_m$, and $\mathcal{J}^0_{\mathbf{F}_p}$ is the identity component of $\mathcal{J}_{\mathbf{F}_p}$. There is a concrete description of $\mathcal{T}$ and of the maps in the exact sequence. Each object in the sequence is equipped with a functorial action of the Hecke algebra $\mathbf{T}$, and the sequence is $\mathbf{T}$-invariant. The p-old quotient $\overline{\mathbf{T}}$ can be viewed as coming from the action of $\mathbf{T}$ on $J_1(M)_{\mathbf{F}_p} \times J_1(M)_{\mathbf{F}_p}$.

By a generalization of [**104**, Lem. 2], the reduction map $J(\overline{\mathbf{Q}}_p)[\ell]^{I_p} \to \mathcal{J}_{\mathbf{F}_p}(\overline{\mathbf{F}}_p)$ is injective. Thus $V = \mathcal{V}_{\mathbf{F}_p}(\overline{\mathbf{F}}_p) \subset \mathcal{J}_{\mathbf{F}_p}(\overline{\mathbf{F}}_p)$. The component group $\Phi = \mathcal{J}_{\mathbf{F}_p}/\mathcal{J}^0_{\mathbf{F}_p}$ is *Eisenstein*, in the sense that it does not contain irreducible representations arising from eigenforms. Since V is irreducible, as a Galois module Φ does not contain an

isomorphic copy of V, so $\mathcal{V}_{\mathbf{F}_p} \subset \mathcal{J}^0_{\mathbf{F}_p}$ and we have the following diagram:



Since $\mathfrak{m}$ acts as 0 on V, the image $\overline{\mathfrak{m}}$ of $\mathfrak{m}$ acts as 0 on the image of V in $J_1(M)_{\mathbf{F}_p} \times J_1(M)_{\mathbf{F}_p}$. If $\overline{\mathfrak{m}} \neq (1)$ then we can optimize the level, so assume $\overline{\mathfrak{m}} = (1)$. Then the image of V in $J_1(M)_{\mathbf{F}_p} \times J_1(M)_{\mathbf{F}_p}$ is 0, so $V_{\mathbf{F}_p} \hookrightarrow \mathcal{T}$.

Let $\mathcal{X}_p(J) := \mathrm{Hom}(\mathcal{T}, \mathbf{G}_m)$ be the *character group* of $\mathcal{T}$. The action of $\mathbf{T}$ on $\mathcal{T}$ induces an action of $\mathbf{T}$ on $\mathcal{X}_p(J)$. Furthermore, $\mathcal{X}_p(J)$ supports an action of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ which, because tori split over a quadratic extension, factors through the Galois group of $\mathbf{F}_{p^2}$. View the Galois action as an action of $\mathrm{Frob}_p \in D_p = \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. With our conventions, the action of Frobenius on the torus is as follows (cf. [**26**, pg. 31]).

Lemma 3.17. *The Frobenius* Frob_p *acts as* pT_p *on* $\mathcal{T}(\overline{\mathbf{F}}_p)$.

Make the identification $\mathcal{T} \cong \mathrm{Hom}_{\mathbf{Z}}(\mathcal{X}_p(J), \mathbf{G}_m)$, so that

$$V \subset \mathcal{T}(\overline{\mathbf{F}}_p)[\ell] = \mathrm{Hom}_{\mathbf{Z}}(\mathcal{X}_p(J), \boldsymbol{\mu}_\ell).$$

By Lemma 3.17, Frob_p acts on $V \subset \mathcal{T}(\overline{\mathbf{F}}_p)$ as $pa_p \in \mathbf{T}/\mathfrak{m}$, i.e., as a *scalar*. The determinant of ρ is $\chi\varepsilon$, so we have simulatenously

$$\det(\rho(\mathrm{Frob}_p)) = \begin{cases} p\varepsilon(p) & \text{and} \\ (pa_p)^2. \end{cases}$$

By Lemma 3.15, $a_p^2 = \varepsilon(p)$, so $p^2 \equiv p \pmod{\ell}$. Since $p \neq \ell$, this can only happen if $p \equiv 1 \pmod{\ell}$, which completes the proof.

3.10. Level optimization using a pivot

In this section we discuss an approach to level optimization that does not rely on multiplicity one results. In this approach, we eliminate a prime p from the level by making use of the rational quaternion algebra that is ramified precisely at p and at a second prime q. The latter prime is, in the simplest case, an appropriate prime number at which ρ ramifies; in more complicated cases, it is an "auxiliary" prime at which ρ is unramified. The central role of q in the argument, and the fact that q stays fixed in the level while p is removed, leads us to refer to q as a "pivot."

The following theorem first appeared in [**86**].

Theorem 3.18. *Let* $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ *be an irreducible continuous representation that arises from an eigenform* f *on* $\Gamma_1(K) \cap \Gamma_0(pq)$ *with* p *and* q *distinct primes that do not divide* ℓK. *Make the* ***key assumption*** *that the representation* ρ *is ramified at* q *and unramified at* p. *Then* ρ *arises from a weight-*2 *eigenform on* $\Gamma_1(K) \cap \Gamma_0(q)$.

The case $\ell = 2$ is not excluded from consideration.

Before sketching the proof, we describe a famous application. Edixhoven suggested to the first author that such an approach might be possible in the context of

Fermat's Last Theorem. We associate to a (hypothetical) solution $a^\ell+b^\ell+c^\ell=0$ of the Fermat equation with $\ell>3$ a Galois representation $E[\ell]$ attached to an elliptic curve E. A theorem of Mazur implies that this representation is irreducible; a theorem of Wiles implies that it arises from a modular form. Using Tate's algorithm, we finds that the discriminant of E is $\Delta_E=\frac{(abc)^{2\ell}}{2^8}$, which is a perfect ℓth power away from 2, and that the conductor of E is $N_E=\mathrm{rad}(abc)=\prod_{p|abc}p$. Let $q=2$; then $E[\ell]$ is ramified at q because $\ell\nmid \mathrm{ord}_2(\Delta_E)=-8$ (see Theorem 2.10), but $E[\ell]$ is unramified at all other primes p, again by Theorem 2.10. To complete the proof of Fermat Last Theorem, we use $q=2$ as a pivot and inductively remove each odd factor from N. One complication that may arise (the second case of Fermat Last Theorem) is that $\ell\mid N$. Upon removing ℓ from the level (using Section 3.1), the weight may initially go up to $\ell+1$. If this occurs, since $k(\rho)=2$ we can use [**32**] to optimize the weight back to 2.

As demonstrated by the application to Fermat, in problems of genuine interest the setup of Theorem 3.18 occurs. There are, however, situations in which it does not apply such as the recent applications of level optimization as a key ingredient to a proof of Artin's conjecture for certain icosahedral Galois representations (see [**10**]).

3.10.1. Shimura curves

We cannot avoid considering Shimura curves. Denote by $X(K,pq)$ the modular curve associated to $\Gamma_1(K)\cap\Gamma_0(pq)$ and let $J:=\mathrm{Jac}(X(K,pq))$ be its Jacobian. Likewise, denote by $X^{pq}(K)$ the Shimura curve associated to the quaternion algebra of discriminant pq. The curve $X^{pq}(K)$ is constructed as follows. Let B be an indefinite quaternion algebra over $\mathbf{Q}$ of discriminant pq. (Up to isomorphism, B is unique.) Let $\mathcal{O}$ be an Eichler order (i.e., intersection of two maximal orders) of level K (i.e., reduced discriminant Kpq) in B. Let Γ_∞ be the group of elements of $\mathcal{O}$ with (reduced) norm 1. After fixing an embedding $B\to M(2,\mathbf{R})$ (an embedding exists because B is indefinite), we obtain in particular an embedding $\Gamma_\infty\hookrightarrow \mathrm{SL}(2,\mathbf{R})$ and therefore an action of Γ_∞ on the upper half-plane $\mathfrak{h}$. Let $X^{pq}(K)$ be the standard canonical model, over $\mathbf{Q}$, of the compact Riemann surface $\Gamma_\infty\backslash\mathfrak{h}$, and let $J'=\mathrm{Jac}(X^{pq}(K))$ denote its Jacobian. The curve $X^{pq}(K)$ is furnished with Hecke correspondences T_n for $n\geq 1$. We write T_n for the endomorphism of J induced by the T_n on $X^{pq}(K)$ via Pic functoriality.

Set $J':=\mathrm{Jac}(X^{pq}(K))$ and $J:=\mathrm{Jac}(X(K,pq))$. Work of Eichler, Jacquet-Langlands, and Shimura (see [**36, 51, 106**]) has uncovered a deep correspondence between certain automorphic forms and certain cusp forms. Combining their work with the isogeny theorem of Faltings [**40**], we find (noncanonically!) a map $J'\to J$ with finite kernel.

The pq-new part of J is $J_{pq\text{-new}}:=\ker(J(K,pq)\longrightarrow J(K,p)^2\oplus J(K,q)^2)$ where the map is induced by Albanese functoriality from the four maps

$$X(K,pq)\rightrightarrows X(K,p)\qquad\text{and}\qquad X(K,pq)\rightrightarrows X(K,q).$$

The image of $J'\to J$ is the pq-new part of J.

3.10.2. Character groups

Amazingly, there seems to be no canonical map $J'\to J$ between the Shimura

and classical Jacobians described in the previous section. Surprisingly, there is a canonical relationship between the character groups of J' and J. The Čerednik-Drinfeld theory gives a description of $X^{pq}(K)$ in characteristic p (see [**14, 30**]). Using this we find a canonical $\mathbf{T}$-equivariant exact sequence

$$0 \to \mathcal{X}_p(J') \to \mathcal{X}_q(J) \to \mathcal{X}_q(J'') \to 0 \tag{3.2}$$

where $J'' = \mathrm{Jac}(X(K,q))^2$. This exact sequence relates a character group "in characteristic p" to two character groups "in characteristic q". We are now prepared to prove the theorem.

3.10.3. Proof

Proof of Theorem 3.18. By our key assumption, the representation ρ is ramified at q, so $\mathfrak{m} \subset \mathbf{T}$ is not q-old. We may as well suppose we are in a situation where we can not optimize the level, so we assume that $\mathfrak{m}$ is not p-old either and hope for a contradiction.

Localization is an exact functor, so the localization

$$0 \longrightarrow \mathcal{X}_p(J')_{\mathfrak{m}} \longrightarrow \mathcal{X}_q(J)_{\mathfrak{m}} \longrightarrow \mathcal{X}_q(J'')_{\mathfrak{m}} \longrightarrow 0 \tag{3.3}$$

of (3.2) is also exact. The Hecke algebra $\mathbf{T}$ acts on $\mathcal{X}_q(J'')$ through a quotient $\overline{\mathbf{T}}$. Since $\mathfrak{m}$ is not q-old, the image of $\mathfrak{m}$ in $\overline{\mathbf{T}}$ generates the unit ideal. Therefore $\mathcal{X}_q(J'')_{\mathfrak{m}} = 0$ and we obtain an isomorphism $\mathcal{X}_p(J')_{\mathfrak{m}} \approx \mathcal{X}_q(J)_{\mathfrak{m}}$. If R is a $\mathbf{T}$-module then $R/\mathfrak{m}R = R_{\mathfrak{m}}/\mathfrak{m}R_{\mathfrak{m}}$ so

$$\mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J) \approx \mathcal{X}_p(J')/\mathfrak{m}\mathcal{X}_p(J'). \tag{3.4}$$

Switching p and q and applying the same argument shows that

$$\mathcal{X}_p(J)/\mathfrak{m}\mathcal{X}_p(J) \approx \mathcal{X}_q(J')/\mathfrak{m}\mathcal{X}_q(J'). \tag{3.5}$$

Both (3.4) and (3.5) are isomorphisms of $\mathbf{T}/\mathfrak{m}$-vector spaces.

By [**6**] we have an isomorphism $J[\mathfrak{m}] \approx \bigoplus_{i=1}^{\lambda} V$, with $\lambda > 0$ and $J'[\mathfrak{m}] \approx \bigoplus_{i=1}^{\nu} V$. (It follows from [**51**] that $\nu > 0$, but we will not use this here.) Our hypothesis that V is unramified automatically propagates to all of $J[\mathfrak{m}] \approx \bigoplus_{i=1}^{\lambda} V$. Since V is irreducible and we are assuming that $\mathfrak{m}$ is not p-old, the same argument as in Section 3.9 shows that $J[\mathfrak{m}] \subset \mathcal{T}[\mathfrak{m}]$ where $\mathcal{T}$ is the toric part of $\mathcal{J}_{\mathbf{F}_p}$. This means that $\dim(\mathcal{X}_p(J)/\mathfrak{m}\mathcal{X}_p(J)) \geq 2\lambda$. Using the same argument with J replaced by J' gives that $\dim(\mathcal{X}_p(J')/\mathfrak{m}\mathcal{X}_p(J')) \geq 2\mu$.

As an I_q-module V is an extension of two copies of the trivial character. This follows from results of Langlands [**67**], since ρ is a mod ℓ representation of $G_{\mathbf{Q}}$ associated to some newform f whose level divides pqK and is divisible by q. (The admissible representation of $\mathrm{GL}(2, \mathbf{Q}_q)$ which is associated to f is a special representation.) Because V is ramified at q and there is an unramified line, we see that $\dim(V^{I_q}) = 1$. Thus $\dim J[\mathfrak{m}]^{I_q} = \lambda$; since $q \neq \ell$ and the action of inertia on character groups is trivial, we see that

$$\mathrm{Hom}(\mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J), \boldsymbol{\mu}_\ell) \subset J[\mathfrak{m}]^{I_q},$$

so $\dim \mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J) \leq \lambda$. A similar argument bounds $\dim \mathcal{X}_q(J')/\mathfrak{m}\mathcal{X}_q(J')$. We obtain the following quadruple of inequalities:

$$\begin{aligned}
\dim \mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J) &\leq \lambda, \\
\dim \mathcal{X}_q(J')/\mathfrak{m}\mathcal{X}_q(J') &\leq \mu, \\
\dim \mathcal{X}_p(J)/\mathfrak{m}\mathcal{X}_p(J) &\geq 2\lambda, \\
\dim \mathcal{X}_p(J')/\mathfrak{m}\mathcal{X}_p(J') &\geq 2\mu.
\end{aligned}$$

Combining these with (3.4, 3.5), we find that

$$\begin{aligned}
2\lambda &\leq \dim \mathcal{X}_p(J)/\mathfrak{m}\mathcal{X}_p(J) \\
&= \dim \mathcal{X}_q(J')/\mathfrak{m}\mathcal{X}_q(J') \\
&\leq \mu
\end{aligned}$$

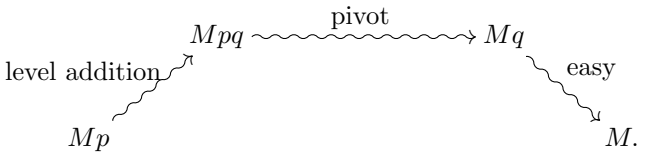
and simulatenously that $2\mu \leq \lambda$. Together these imply that $4\lambda \leq \lambda$ so $\lambda = 0$. But Proposition 3.2 implies that the multiplicity of ρ in $J[\mathfrak{m}]$ is strictly positive. This contradiction implies that our assumption that $\mathfrak{m}$ is not p-old is false, hence $\mathfrak{m}$ is p-old and ρ arises from an eigenform on $\Gamma_1(K) \cap \Gamma_0(q)$. □

3.11. Level optimization with multiplicity one

Theorem 3.19. *Suppose $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ is an irreducible multiplicity one representation that arises from a weight-2 newform f on $\Gamma_1(M) \cap \Gamma_0(p)$ and that p is unramified. Then there is a newform on $\Gamma_1(M)$ that also gives rise to ρ.*

We sketch a proof, under the assumption that $\ell > 2$. Buzzard [**9**] has given a proof when $\ell = 2$; his result has been combined with the results of [**28**] to prove a Wiles-like lifting theorem valid for many representations when $\ell = 2$, and hence (thanks to Taylor) to establish new examples of Artin's conjecture (see [**10**]).

The following diagram illustrates the multiplicity one argument:



The pivot step is potentially the hardest; though it resembles the pivot step of Section 3.10, but the symmetry is broken. In Section 3.10 we knew that q could not be removed from the level, but here q can be.

We manufacture q as follows. Pick q to be one of the (infinitely many) primes not dividing $Mp\ell$ such that the following conditions both hold:

(1) $\begin{cases} \rho(\mathrm{Frob}_q) \text{ is not a scalar, } \textit{or} \\ q \not\equiv 1 \pmod{\ell}. \end{cases}$

(2) $\begin{cases} \text{the ratio of the eigenvalues of} \\ \rho(\mathrm{Frob}_q) \text{ is either } q \text{ or } 1/q. \end{cases}$

The second condition means that the characteristic polynomial of $\rho(\mathrm{Frob}_q)$ is of the form $(x-a)(x-qa)$ for some $a \in \overline{\mathbf{F}}_\ell^*$.

Lemma 3.20. *There are infinitely many primes q that simultaneously satisfy both of the two conditions listed above.*

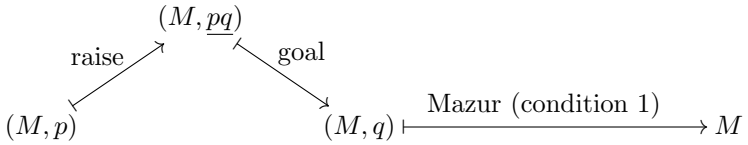
Proof. First assume that $\ell > 2$. Using the Cebotarev density theorem, find infinitely many primes q such $\rho(\mathrm{Frob}_q) = \rho(c)$ where c denotes complex conjugation. The eigenvalues of $\rho(c)$ are 1 and -1 (Exercise 8), so their ratio is -1. This ratio is equal to q because

$$-1 = \chi(c) = \det(\rho(\mathrm{Frob}_q)) = \chi(\mathrm{Frob}_q) \equiv q \pmod{\ell},$$

and $q \not\equiv 1 \pmod{\ell}$ because ℓ is odd.

Next assume that $\ell = 2$. Because ρ is irreducible, the image $\rho(G_{\mathbf{Q}}) \subset \mathrm{GL}(2, \overline{\mathbf{F}}_2)$ has even order. After a possible change of basis we find $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \rho(G_{\mathbf{Q}})$. Using Cebotarev density, we find infinitely many q with $\rho(\mathrm{Frob}_q)$ conjugate to $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. For such q, condition (1) is satisfied. Condition (2) is also satisfied because the ratio of the eigenvalues is 1 which, because q is an odd prime, is congruent to q modulo $\ell = 2$. □

Sketch of proof of Theorem 3.19. Choose q as in Lemma 3.20. With q thus chosen, we can raise the level. More precisely, there exists a pq-new form on $\Gamma_1(M) \cap \Gamma_0(pq)$. We illustrate this as follows.



We underline pq to emphasize that the situation at level (M, pq) is symmetrical in p and q.

Let $J = J(M, pq)$; there is a maximal ideal $\mathfrak{m}$ in $\mathbf{T} = \mathbf{Z}[\ldots T_n \ldots] \subset \mathrm{End}\, J$ attached to the pq-newform f that gives rise to ρ. Applying the multiplicity one hypothesis at level Mpq, we have $J[\mathfrak{m}] = V$ where V is a $\mathbf{T}/\mathfrak{m}$-vector space that supports ρ. In everything so far, M can be divisible by 2; the distinction between whether or not 2 divides M arises mainly in verifying the multiplicity one hypothesis.

Let $J' = J^{pq}(M)$ be the Shimura curve analogue of $J_1(M)$. As described in Section 3.10, J' is constructed in a similar manner as $J_1(M)$, but with $M_2(\mathbf{Q})$ replaced by a quaternion algebra. Of primary importance is that $J'[\mathfrak{m}] \approx \bigoplus_{i=1}^{\nu} V$, for some $\nu \geq 1$. This follows *morally* because ρ arises from a pq-new form, though the actual argument is quite involved.

Assume that we cannot optimize the level. We have an exact sequence of character groups

$$0 \longrightarrow \mathcal{X}_p(J') \longrightarrow \mathcal{X}_q(J) \longrightarrow \mathcal{X}_q(J(M,q)^2) \longrightarrow 0.$$

After localizing at $\mathfrak{m}$ as in (3.3), we discover that

$$\dim_{\mathbf{T}/\mathfrak{m}} \mathcal{X}_p(J')/\mathfrak{m}\mathcal{X}_p(J') = \dim_{\mathbf{T}/\mathfrak{m}} \mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J). \tag{3.6}$$

Furthermore, since the component group of J' at p is a quotient of $\mathcal{X}_q(J(M,q)^2)$, we find that $V \hookrightarrow (J'[\mathfrak{m}]^{I_p})^{\mathrm{toric}}$. Thus $\dim \mathcal{X}_p(J')/\mathfrak{m}\mathcal{X}_p(J') \geq 2$, so (3.6) implies that $\dim \mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J) \geq 2$. The endomorphism Frob_q acts as a scalar (cf. Lemma 3.17) on

$$J[\mathfrak{m}]^{\mathrm{toric}} = \mathrm{Hom}(\mathcal{X}_q(J)/\mathfrak{m}\mathcal{X}_q(J), \mu_\ell).$$

Furthermore, $J[\mathfrak{m}]^{\mathrm{toric}} \subset J[\mathfrak{m}]$ and both $J[\mathfrak{m}]^{\mathrm{toric}}$ and $J[\mathfrak{m}]$ have dimension 2, so Frob_q acts as a scalar on $J[\mathfrak{m}]$. If $q \not\equiv 1 \pmod{\ell}$ then we could use Mazur's principle to optimize the level, so by condition 1 we may assume that $\rho(\mathrm{Frob}_q)$ is not a scalar. This contradiction completes the sketch of the proof. $\square$

CHAPTER 4

Exercises

The following exercises were used in the Park City problem sessions. D. Savitt, K. Kedlaya, and B. Conrad contributed some of the problems. In Section 4.2, we provide several solutions, many of which were suggested by students in the problem sessions. The solution of some of the problems in this section requires facts beyond those stated explicitly in this paper.

4.1. Exercises

Exercise 1. Suppose $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}_\ell^*$ is a one-dimensional continuous odd Galois representation.

(1) Give an example to show that ρ need not be a power of the mod ℓ cyclotomic character.
(2) Assume that ρ is unramified outside ℓ. Deduce that ρ is a power of the mod ℓ cyclotomic character.

Exercise 2. The principal congruence subgroup $\Gamma(N)$ of level N is the kernel of the reduction map $\mathrm{SL}(2,\mathbf{Z}) \to \mathrm{SL}(2,\mathbf{Z}/N\mathbf{Z})$. The subgroup $\Gamma_1(N)$ consists of matrices of the form $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ modulo N. Let $\Gamma \subset \mathrm{SL}(2,\mathbf{Z})$ be a subgroup that contains $\Gamma(N)$ for some N. Show that there exists $g \in \mathrm{GL}(2,\mathbf{Q})$ such that the conjugate $g^{-1}\Gamma g$, which is a subgroup of $\mathrm{GL}(2,\mathbf{Q})$, contains $\Gamma_1(N^2)$.

Exercise 3. Let k be a finite field of characteristic greater than 2, and consider an odd representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2,k)$. Prove that ρ is irreducible if and only if ρ is absolutely irreducible. (A representation is absolutely irreducible if it remains irreducible after composing with the embedding $\mathrm{GL}(2,\mathbf{F}_\ell) \hookrightarrow \mathrm{GL}(2,\overline{\mathbf{F}}_\ell)$.) Give an example to show that this assertion is false when k has characteristic 2.

Exercise 4. Let $A/\mathbf{Q}$ be an elliptic curve. Show that the group of $\mathbf{Q}$-rational endomorphisms $\mathrm{End}(A)$ of A is equal to $\mathbf{Z}$; that is, integer multiplications are the only $\mathbf{Q}$-rational endomorphisms of A. Assume further that A is isolated in its isogeny class, in the sense that if B is an elliptic curve that is isogenous to A over $\mathbf{Q}$, then A and B are isomorphic over $\mathbf{Q}$. Show that, for every prime number ℓ, the representation

$$\rho_\ell : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(A[\ell]) \approx \mathrm{GL}(2,\mathbf{F}_\ell)$$

is irreducible. Must ρ_ℓ be absolutely irreducible?

Exercise 5. Let $A/\mathbf{Q}$ be an elliptic curve and assume that for all ℓ the representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(A[\ell])$ is irreducible. Deduce that A is isolated in its isogeny class. This is the converse of Exercise 4.

Exercise 6. Suppose $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, \mathbf{F}_\ell)$ arises from the ℓ-torsion of an elliptic curve. Verify, using standard properties of the Weil pairing, that $\det(\rho)$ is the mod ℓ cyclotomic character.

Exercise 7. Let $f \in S_k(\Gamma_1(N))$ be a modular form that is an eigenform for all the Hecke operators T_p and for the diamond bracket operators $\langle d \rangle$. Let

$$\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \mathbf{C}^*$$

be the character of f, so $\langle d \rangle f = \varepsilon(d) f$ for all $d \in (\mathbf{Z}/N\mathbf{Z})^*$.

(1) Show that f satisfies the following equation:
for any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$,

$$f(z) = \varepsilon(d)(cz + d)^{-k} f\Big(\frac{az + b}{cz + d}\Big).$$

(2) Conclude that $\varepsilon(-1) = (-1)^k$.
(3) Choose a prime ℓ and let ρ be one of the mod ℓ Galois representations associated to f. We have $\det(\rho) = \varepsilon \cdot \chi^{k-1}$ where χ is the mod ℓ cyclotomic character. Deduce that ρ is odd, in the sense that $\det(\rho(c)) = -1$ for c complex conjugation.

Exercise 8. Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \overline{\mathbf{F}}_\ell)$ be an odd Galois representation, and let $c \in G_{\mathbf{Q}}$ denote complex conjugation.

(1) Prove that if $\ell \neq 2$ then $\rho(c)$ is conjugate over $\overline{\mathbf{F}}_\ell$ to the matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.
(2) Give an example to show that when $\ell = 2$, the matrix $\rho(c)$ need not be conjugate to $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Exercise 9. Show that there exists a *non-continuous* homomorphism

$$\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \{\pm 1\}$$

where $\{\pm 1\}$ has the discrete topology; equivalently, that there is a non-closed subgroup of index two in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. To accomplish this, you must produce a map $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \{\pm 1\}$ such that

(1) ρ is a homomorphism, and
(2) ρ does not factor through $\mathrm{Gal}(K/\mathbf{Q})$ for any *finite* Galois extension $K/\mathbf{Q}$.

Exercise 10. A potential difficulty is that a representation ρ arising from a modular form sometimes takes values in a slightly smaller field than $\mathcal{O}/\lambda$. For example, let f be one of the two conjugate normalized eigenforms in $S_2(\Gamma_0(23))$. Then

$$f = q + \alpha q^2 + (-2\alpha - 1)q^3 + (-\alpha - 1)q^4 + 2\alpha q^5 + \cdots$$

with $\alpha^2 + \alpha - 1 = 0$. The coefficients of f lie in $\mathcal{O} = \mathbf{Z}[\alpha] = \mathbf{Z}[\frac{1+\sqrt{5}}{2}]$. Take λ to be the unique prime of $\mathcal{O}$ lying over 2; then $\mathcal{O}/\lambda \cong \mathbf{F}_4$, so $\overline{\rho}_{f,\lambda}$ is a homomorphism into $\mathrm{GL}(2, \mathbf{F}_4)$. Show that if $p \neq 2$ then $a_p \in \mathbf{Z}[\sqrt{5}]$, so that $\overline{\rho}_{f,\lambda}$ possesses a model over $\mathrm{GL}(2, \mathbf{F}_2)$.

Exercise 11. Let $A/\mathbf{Q}$ be an elliptic curve and $\ell \neq 2$ be a prime.

(1) Prove that the field $\mathbf{Q}(A[\ell])$ generated by the coordinates of the points in $A[\ell]$ is strictly larger than $\mathbf{Q}$.
(2) Given an example of an elliptic curve A such that $\mathbf{Q}(A[2]) = \mathbf{Q}$.

Exercise 12. Let A be an elliptic curve over $\mathbf{Q}$ defined by a Weierstrass equation $y^2 = x^3 + ax + b$ with $a, b \in \mathbf{Q}$.

(1) Describe the Galois representation
$$\rho = \rho_{A,2} : G_{\mathbf{Q}} \to \mathrm{GL}(2, \mathbf{F}_2).$$
(2) Give necessary and sufficient conditions for ρ to be reducible.
(3) Choose a prime p, and give an example in which ρ is ramified only at p.

Exercise 13. Let ε and ρ be a pair of continuous homomorphisms from $G_{\mathbf{Q}}$ to $\mathbf{F}_\ell^*$. Suppose that for all primes p at which both ε and ρ are unramified we have
$$\rho(\mathrm{Frob}_p) = \varepsilon(\mathrm{Frob}_p)p^i \in \mathbf{F}_\ell.$$
Deduce that $\rho = \varepsilon \cdot \chi^i$ where χ is the mod ℓ cyclotomic character.

Exercise 14. Let $A/\mathbf{Q}$ be an elliptic curve of conductor N, and let p be a prime number not dividing N. Denote by $\tilde{A}$ the mod p reduction of A. The Frobenius endomorphism $\Phi = \Phi_p : \tilde{A} \to \tilde{A}$ sends an affine point (x, y) to (x^p, y^p) and fixes ∞. The characteristic polynomial of the endomorphism induced by Φ on the Tate module of $\tilde{A}$ at some (any) prime $\ell \neq p$ is $X^2 - \mathrm{tr}(\Phi)X + \deg(\Phi)$.

(1) Show that $\deg(\Phi) = p$.
(2) Show that $\mathrm{tr}(\Phi) = p + 1 - \#A(\mathbf{F}_p)$, that is, "$\mathrm{tr}(\Phi) = a_p$."
(3) Choose a prime $\ell \nmid pN$. Then $\tilde{A}[\ell]$ is a vector space of dimension two over $\mathbf{F}_\ell$, and Φ induces a map $\tilde{A}[\ell] \to \tilde{A}[\ell]$. Show that this is the same as the map induced by some choice of $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.
(4) Conclude that
$$\mathrm{tr}(\rho_{A,\ell}(\mathrm{Frob}_p)) = p + 1 - \#A(F_p) \pmod{\ell}.$$

Exercise 15. Let $A/\mathbf{Q}$ be an elliptic curve of conductor N, and let ℓ be a prime. Show that any prime p not dividing ℓN is unramified in $\mathbf{Q}(A[\ell])$. You may use the following fact which is proved using formal groups (see, e.g., [**109**, Prop. 3.1]): **Fact:** The map $A[\ell] \to \tilde{A}(\overline{\mathbf{F}}_p)$ is injective, where $\tilde{A}$ is the reduction of A modulo p.

Exercise 16. Show that the fundamental character of level 1 is the cyclotomic character $\chi|_{I_t}$. (Hint: This is trickier than it first appears, and requires Wilson's theorem from elementary number theory.)

Exercise 17. For each of the following semistable elliptic curves A, and each ℓ at which $\rho_{A,\ell}$ is *irreducible*, use Theorem 2.10 to compute Serre's minimal weight $k(\rho_{A,\ell})$ and level $N(\rho_{A,\ell})$.

N	$\lvert\Delta\rvert$	reducible ℓ	A
30	$2^4 \cdot 3^5 \cdot 5$	2, 3	$y^2 + xy + y = x^3 + x + 2$
210	$2^{12} \cdot 3^3 \cdot 5 \cdot 7$	2, 3	$y^2 + xy = x^3 - 41x - 39$
330	$2^4 \cdot 3^2 \cdot 5^4 \cdot 11^2$	2	$y^2 + xy = x^3 + x^2 - 102x + 324$
455	$5^3 \cdot 7^4 \cdot 13$	2	$y^2 + xy = x^3 - x^2 - 50x + 111$
2926	$2^8 \cdot 7^3 \cdot 11^4 \cdot 19^2$	2	$y^2 + xy + y = x^3 - x^2 + 1934x - 1935$

Attempt to verify Serre's conjecture directly in some of these cases.

Exercise 18. Let M be a positive integer and let p be a prime. Show that there is an injective linear map
$$S_2(\Gamma_1(M)) \hookrightarrow S_2(\Gamma_1(pM))$$
sending $f(q)$ to $f(q^p)$.

Exercise 19. Let M be an integer such that $S_2(\Gamma_1(M))$ has positive dimension, and let p be a prime (thus $M = 11$ or $M \geq 13$).

(1) Let $f \in S_2(\Gamma_1(M))$ be an eigenvector for T_p with eigenvalue λ. Show that T_p acting on $S_2(\Gamma_1(Mp))$ preserves the two-dimensional subspace generated by f and $f(pz)$ (see Section 1.5 for the definition of T_p when p divides the level). Show furthermore that if $\lambda^2 \neq 4p$ then T_p is diagonalizable on this 2-dimensional space. What are the eigenvalues of T_p on this space? In fact, one never has $\lambda^2 = 4p$; see [**16**] for more details.
(2) Show that for any $r > 2$, the Hecke operator T_p on $S_2(\Gamma_1(Mp^r))$ is not diagonalizable.
(3) Deduce that for $r > 2$ the Hecke algebra $\mathbf{T}$ associated to $S_2(\Gamma_1(Mp^r))$ has nilpotent elements, so it is not an order in a product of rings of integers of number fields.

Exercise 20. Let N be a positive integer. Show that the Hecke algebra $\mathbf{T} = \mathbf{Z}[\ldots T_n \ldots] \subset \mathrm{End}(J_1(N))$ is of finite rank as a $\mathbf{Z}$-module.

Exercise 21. Suppose $N = pM$ with $(p, M) = 1$. There is an injection

$$S_2(\Gamma_1(M)) \oplus S_2(\Gamma_1(M)) \hookrightarrow S_2(\Gamma_1(M) \cap \Gamma_0(p))$$

given by $(f, g) \mapsto f(q) + g(q^p)$. The Hecke algebra $\mathbf{T} = \mathbf{T}_N$ acts through a quotient $\overline{\mathbf{T}}$ on the image of $S_2(\Gamma_1(M)) \oplus S_2(\Gamma_1(M))$. Suppose $\mathfrak{m} \subset \mathbf{T}$ is a maximal ideal that arises by pullback from a maximal ideal in $\overline{\mathbf{T}}$. Show that $\rho_{\mathfrak{m}}$ arises from a modular form of level M.

4.2. Solutions

Solution 1. 1. Let p be a prime different from ℓ and let

$$\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gal}(\mathbf{Q}(\sqrt{p})/\mathbf{Q}) \approx \{\pm 1\} \hookrightarrow \mathbf{F}_\ell^*.$$

2. Let $K = \overline{\mathbf{Q}}^{\ker(\rho)}$. Then $K/\mathbf{Q}$ is abelian and ramified only at ℓ, so $K \subset \mathbf{Q}(\zeta_{\ell^\infty})$. But $[K : \mathbf{Q}] \mid \ell - 1$ so $K \subset \mathbf{Q}(\zeta_\ell)$.

Solution 2. Conjugate using $g = \left(\begin{smallmatrix} N & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Solution 3. If ρ is absolutely irreducible then it is irreducible, so assume that ρ is irreducible. If ρ is reducible over the algebraic closure $\overline{k}$ of k, then there is a vector $v \in \overline{k}^{\oplus 2}$ that generates a one-dimensional subspace stable under ρ. In particular, v is stable under complex conjugation, which has characteristic polynomial $x^2 - 1 = (x - 1)(x + 1)$. Since $-1 \neq 1$, this means that v must lie in one of the two 1-dimensional eigenspaces of complex conjugation, so v is a scalar multiple of an element w of $k^{\oplus 2}$. Then ρ leaves the subspace of $k^{\oplus 2}$ spanned by w invariant, so ρ is reducible, which contradicts our assumption.

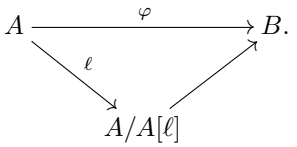
Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}(2, \mathbf{F}_2)$ be any continuous representation whose image is the subgroup generated by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. Then ρ is irreducible because it has no one-dimensional invariant subspaces over $\mathbf{F}_2$. However, the matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ is diagonalizable over $\mathbf{F}_4$.

Solution 4. Suppose $\varphi \in \mathrm{End}(E)$ is a nonzero endomorphism. The induced map $d\varphi$ on the differentials $H^0(A,\Omega) \approx \mathbf{Q}$ is multiplication by an integer n, so $d(\varphi - n) = 0$ which implies that $\varphi = n$.

Suppose that ρ_ℓ is reducible, so that there is a one-dimensional Galois stable subspace $V \subset A[\ell]$. The quotient $B = A/V$ is then an elliptic curve over $\mathbf{Q}$ and there is an isogeny $\pi : A \to B$ of degree ℓ. Because A is isolated in its isogeny class we have that $B = A$, so there is an endomorphism of A of degree ℓ. But all $\mathbf{Q}$-rational endomorphisms are multiplication by an integer, and multiplication by an integer has degree a perfect square.

The elliptic curve E given by the equation $y^2 = x^3 - 7x - 7$ has the property that $E[2]$ is irreducible but not absolutely irreducible. To see this, note that the splitting field of $x^3 - 7x - 7$ has Galois group cyclic of order 3.

Solution 5. Suppose all $\rho_{A,\ell}$ are irreducible, yet there exists an isogeny $\varphi : A \to B$ with $B \not\approx A$. Choose φ to have minimal possible degree and let $d = \deg(\varphi) > 1$. Let ℓ be the smallest prime divisor of d and choose a point $x \in \ker(\varphi)$ of exact order ℓ. If the order-ℓ cyclic subgroup generated by x is Galois stable, then $\rho_{A,\ell}$ is reducible, which is contrary to our assumption. Thus $\ker(\varphi)$ contains the full ℓ-torsion subgroup $A[\ell]$ of A. In particular, φ factors as illustrated below:



Since $A/A[\ell] \cong A$, there is an isogeny from A to B of degree equal to d/ℓ^2, which contradicts our assumption that d is minimal.

Solution 6. The Weil pairing $(\ ,\ ) : A[\ell] \times A[\ell] \to \boldsymbol{\mu}_\ell$ can be viewed as a map

$$\bigwedge^2 A[\ell] \xrightarrow{\cong} \boldsymbol{\mu}_\ell$$

sending $P \wedge Q$ to (P,Q). For any $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, we have $(P^\sigma, Q^\sigma) = (P,Q)^\sigma$. With the action $(P \wedge Q)^\sigma = P^\sigma \wedge Q^\sigma$, the map $\bigwedge^2 A[\ell] \to \boldsymbol{\mu}_\ell$ is a map of Galois modules. To compute $\det(\rho(\sigma))$ observe that if e_1, e_2 is a basis for $A[\ell]$, and $\rho(\sigma) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, then

$$\begin{aligned} \sigma(e_1 \wedge e_2) &= (ae_1 + ce_2) \wedge (be_1 + de_2) \\ &= (ad - bc)e_1 \wedge e_2 = \det(\rho(\sigma))e_1 \wedge e_2 \end{aligned}$$

Thus $\bigwedge^2 A[\ell]$ gives the one-dimensional representation $\det(\rho)$. Since $\bigwedge^2 A[\ell]$ is isomorphic to $\boldsymbol{\mu}_\ell$ it follows that $\det(\rho) = \chi$.

Solution 7. The definition of $\langle d \rangle$ is as follows: choose any matrix $\sigma_d \in \Gamma_0(N)$ such that $\sigma_d \equiv \left(\begin{smallmatrix} d & 0 \\ 0 & d^{-1} \end{smallmatrix}\right) \pmod N$; then $\langle d \rangle f = f|_{\sigma_d}$. Observe that $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$ and the matrices σ_d with $(d,N) = 1$ and $d < N$ are a system of coset representatives. Thus any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ can be written in the form $\sigma_d \cdot g$ for some $g \in \Gamma_1(N)$. We have

$$f = f|_{\sigma_d g} = (f|_{\sigma_d})|_g = (\varepsilon(d)f)|_g = \varepsilon(d)f|_g = \varepsilon(d)(cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right).$$

Solution 8.

(1) Since $c^2 = 1$, the minimal polynomial f of $\rho(c)$ divides $x^2 - 1$. Thus f is either $x+1$, $x-1$, or x^2-1. If $f = x+1$ then $\rho(c) = -1 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. This implies that $\det(\rho(c)) = (-1)^2 = 1$, which is a contradiction since $\det(\rho(c)) = -1$ and the characteristic of the base field is odd. If $f = x-1$, then $\rho(c) = 1$; again a contradiction. Thus the minimal polynomial of $\rho(c)$ is $x^2 - 1 = (x-1)(x+1)$. Since $-1 \neq 1$ there is a basis of eigenvectors for $\rho(c)$ such that the matrix of $\rho(c)$ with respect to this basis is $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

(2) The following example shows that when $\ell = 2$ the matrix of $\rho_{A,\ell}$ need not be conjugate to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Let A be the elliptic curve over $\mathbf{Q}$ defined by $y^2 = x(x^2 - a)$ with $a \in \mathbf{Q}$ not square. Then

$$A[2] = \{\infty, (0,0), (\sqrt{a}, 0), (-\sqrt{a}, 0)\}.$$

The action of c on the basis $(0,0), (-\sqrt{a}, 0)$ is represented by the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, since $c(-\sqrt{a}, 0) = (\sqrt{a}, 0) = (0,0) + (-\sqrt{a}, 0)$.

Solution 9. The extension $\mathbf{Q}(\sqrt{d}, d \in \mathbf{Q}^*/(\mathbf{Q}^*)^2)$ is an extension of $\mathbf{Q}$ with Galois group $X \approx \prod \mathbf{F}_2$. The index-two open subgroups of X correspond to the quadratic extensions of $\mathbf{Q}$. However, Zorn's lemma implies that X contains many more index-two subgroups, which can be seen more precisely as follows.

(1) Choose a sequence $p_1, p_2, p_3, \ldots$ of distinct prime numbers. Define $\rho_1 : G_{\mathbf{Q}} \to \prod \mathbf{F}_2$ by

$$\rho_1(\sigma)_i = \begin{cases} 0 & \text{if } \sigma \text{ acts trivially on } \mathbf{Q}(\sqrt{p_i}), \\ 1 & \text{otherwise} \end{cases}$$

Thus ρ_1 is just

$$G_{\mathbf{Q}} \to \mathrm{Gal}(\mathbf{Q}(\sqrt{p_1}, \sqrt{p_2}, \ldots)/\mathbf{Q}) \approx \prod \mathbf{F}_2.$$

(2) Let $\oplus \mathbf{F}_2 \subset \prod \mathbf{F}_2$ be the subgroup of elements having only finitely many nonzero coordinates. Then $\prod \mathbf{F}_2 / \oplus \mathbf{F}_2$ is a vector space over $\mathbf{F}_2$ of dimension > 0. By Zorn's lemma, there is a basis $\mathcal{B}$ of $\prod \mathbf{F}_2 / \oplus \mathbf{F}_2$. Let $b \in \mathcal{B}$ and let W be the subspace spanned by $\mathcal{B} - \{b\}$. Then $V = (\prod \mathbf{F}_2 / \oplus \mathbf{F}_2)/W$ is an $\mathbf{F}_2$-vector space of dimensional 1.

(3) Let ρ be the composite map

$$\begin{array}{ccccc} \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) & & & & \\ \downarrow & \searrow^{\rho} & & & \\ \prod \mathbf{F}_2 & \longrightarrow & V & \xrightarrow[\cong]{} & \{\pm 1\} \end{array}$$

(4) Let $H = \ker(\rho) \subset \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. If $\sigma(\sqrt{p_i}) = -\sqrt{p_i}$ and $\sigma(\sqrt{p_j}) = \sqrt{p_j}$ for $i \neq j$, then $\sigma \in H$. Thus H does not fix any $\mathbf{Q}(\sqrt{p_i})$, so the fixed field of H equals $\mathbf{Q}$. The largest finite Galois group quotient through which ρ factors is then $\mathrm{Gal}(\mathbf{Q}/\mathbf{Q}) = \{1\}$. Since $\rho \neq 1$, we conclude that ρ does not factor through any finite Galois group quotient, which proves that ρ is not continuous.

Solution 10. We have $f = f_1 + \alpha f_2$ with $f_1, f_2 \in S_2(\Gamma_0(23))$ and

$$f_1 = q - q^3 - q^4 + \cdots$$
$$f_2 = q^2 - 2q^3 - q^4 + 2q^5 + \cdots .$$

Because $S_2(\Gamma_0(23))$ has dimension 2, it is spanned by f_1 and f_2. Let $\eta(q) = q^{\frac{1}{24}} \prod_{n\geq 1}(1-q^n)$. Then $g = (\eta(q)\eta(q^{23}))^2 \in S_2(\Gamma_0(23))$. Expanding we find that $g = q^2 - 2q^3 + \cdots$, so $g = f_2$. Next observe that g is a power series in q^2 modulo 2:

$$\begin{aligned} g &= q^2 \prod (1-q^n)^2(1-q^{23n})^2 \\ &\equiv q^2 \prod (1-q^{2n})(1-q^{46n}) \pmod 2 \\ &\equiv q^2 \prod (1+q^{2n}+q^{46n}+q^{48n}) \pmod 2 \end{aligned}$$

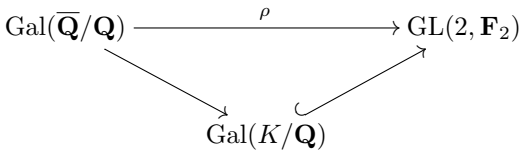
Thus the coefficient in f_2 of q^p with $p \neq 2$ prime is even, and the proposition follows.

Solution 11.

(1) Let $\zeta \in \boldsymbol{\mu}_\ell$ be a primitive ℓth root of unity. Since $\bigwedge^2 A[\ell] \cong \boldsymbol{\mu}_\ell$, there exists $P, Q \in A[\ell]$ such that $P \wedge Q = \zeta$. Since $\ell > 2$ there exists σ such that $\zeta^\sigma \neq \zeta$, hence $P^\sigma \wedge Q^\sigma \neq P \wedge Q$. This is impossible if all ℓ-torsion is rational, since then $P^\sigma = P$ and $Q^\sigma = Q$.
(2) Consider the elliptic curve defined by $y^2 = (x-a)(x-b)(x-c)$ where a, b, c are distinct rational numbers.

Solution 12.

(1) Let K be the splitting field of $x^3 + ax + b$. Then ρ embeds $\mathrm{Gal}(K/\mathbf{Q})$ in $\mathrm{GL}(2, \mathbf{F}_2)$:



(2) The representation ρ is reducible exactly when the polynomial $x^3 + ax + b$ has a rational root.
(3) Examples: $y^2 = x(x^2 - 23)$, $y^2 = x^3 + x - 1$.

Solution 13. Consider the character $\tau = \varepsilon\chi/\rho$. By assumption, $\tau(\mathrm{Frob}_p) = 1$ for all unramified p. Let K be an extension of $\mathbf{Q}$ such that τ factors through $\mathrm{Gal}(K/\mathbf{Q})$. For any $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$, the Cebotarev density theorem implies that there are infinitely many primes p such that $\mathrm{Frob}_p = \sigma$. Thus for any σ, $\tau(\sigma) = \tau(\mathrm{Frob}_p) = 1$, so $\tau = 1$ and hence $\rho = \varepsilon\chi$.

Solution 14.

(1) See, e.g., [**109**, 2.11].
(2) By [**109**, 5.5], $\Phi - 1$ is separable, so $\#A(\mathbf{F}_p) = \deg(\Phi - 1)$. Since Φ has degree p, there exists an isogeny $\overline{\Phi}$ (the dual isogeny, see [**109**, III.6]),

such that $\Phi\overline{\Phi} = p$. Letting bars denote the dual isogeny, we have

$$\begin{aligned}\#A(\mathbf{F}_p) = \deg(\Phi - 1) &= (\Phi - 1)\overline{(\Phi - 1)} \\ &= \Phi\overline{\Phi} - \Phi - \overline{\Phi} + 1 \\ &= p - \mathrm{tr}(\Phi) + 1\end{aligned}$$

(3) Both maps are pth powering on coordinates.

Solution 15. Since $\ell \neq p$ and A has good reduction at p, the natural map $A[\ell] \to \tilde{A}[\ell]$ is an isomorphism. We have the following commutative diagram

$$\begin{array}{ccc} \mathrm{Gal}(\mathbf{Q}_p(A[\ell])/\mathbf{Q}_p) & \hookrightarrow & \mathrm{Aut}(A[\ell]) \\ \downarrow & & \downarrow \cong \\ \mathrm{Gal}((\mathcal{O}/\lambda)_{\mathbf{F}_p}) & \longrightarrow & \mathrm{Aut}(\tilde{A}[\ell]) \end{array}$$

It follows that the first vertical map must be injective, which is the same as $\mathbf{Q}_p(A[\ell])$ being unramified over $\mathbf{Q}_p$.

Solution 16. The fundamental character Ψ of level one is the composition

$$\mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[\ell-1]{\ell})/\mathbf{Q}_\ell^{\mathrm{nr}}) \to \boldsymbol{\mu}_{\ell-1}(\overline{\mathbf{Q}}_\ell) \to \boldsymbol{\mu}_{\ell-1}(\overline{\mathbf{F}}_\ell^*) = \mathbf{F}_\ell^*.$$

Let π be such that $\pi^{\ell-1} = \ell$. Then $\Psi(\sigma) = \frac{\sigma(\pi)}{\pi} \pmod{\pi}$. Let $\zeta \in \overline{\mathbf{Q}}_\ell$ be a primitive ℓth root of unity. Now

$$\prod_{a=1}^{\ell-1}(\zeta^a - 1) = \ell,$$

so

$$(\zeta - 1)^{\ell-1} \prod_{a=1}^{\ell-1}, \frac{\zeta^a - 1}{\zeta - 1} = \ell$$

and (this is where Wilson's theorem is used),

$$\prod_{a=1}^{\ell-1} \frac{\zeta_\ell^a - 1}{\zeta - 1} \equiv 1 \pmod{\zeta - 1}.$$

Since the polynomial $x^{\ell-1} - 1$ has roots over $\mathbf{F}_\ell$, by Hensel's lemma there is a unit $u \in \mathbf{Q}_\ell(\pi)$ such that

$$u^{\ell-1} = \prod_{a=1}^{\ell-1} \frac{\zeta^a - 1}{\zeta - 1}.$$

We can take $\pi = (\zeta - 1)u$. Then

$$\begin{aligned}\frac{\sigma(\pi)}{\pi} &= \frac{(\zeta^{\chi(\sigma)} - 1)\sigma(u)}{(\zeta - 1)u} \\ &= \frac{(\zeta - 1)(\zeta^{\chi(\sigma)-1} + \cdots + 1)\sigma(u)}{(\zeta - 1)u} \\ &= (\zeta^{\chi(\sigma)-1} + \cdots + 1)\sigma(u)/u \\ &\equiv \chi(\sigma) \pmod{\zeta - 1}.\end{aligned}$$

Solution 17. We write $N = N(\rho)$ and $k = k(\rho)$ to save space. The essential tool is Theorem 2.10.

(1) $\ell = 5$: $N = 6$, $k = 6$, $\ell > 5$, $N = 30$, $k = 2$.
(2) $\ell = 5$: $N = 2 \cdot 3 \cdot 7$, $k = 6$; $\ell = 7$: $N = 2 \cdot 3 \cdot 5$, $k = 8$; $\ell > 7$: $N = 2 \cdot 3 \cdot 5 \cdot 7$, $k = 2$.
(3) $\ell = 3$: $N = 2 \cdot 5 \cdot 11$, $k = 4$; $\ell = 5$: $N = 2 \cdot 3 \cdot 11$, $k = 6$; $\ell = 7$: $N = 2 \cdot 3 \cdot 5 \cdot 11$, $k = 2$; $\ell = 11$: $N = 2 \cdot 3 \cdot 5$, $k = 12$; $\ell > 11$: $N = 2 \cdot 3 \cdot 5 \cdot 11$, $k = 2$.
(4) $\ell = 3$: $N = 7 \cdot 13$, $k = 2$; $\ell = 5$: $N = 7 \cdot 13$, $k = 6$; $\ell = 7$: $N = 5 \cdot 13$, $k = 8$; $\ell = 13$: $N = 5 \cdot 7$, $k = 14$; $\ell = 11, \ell > 13$: $N = 5 \cdot 7 \cdot 13$, $k = 2$.
(5) $\ell = 3$: $N = 2 \cdot 11 \cdot 19$, $k = 2$; $\ell = 7$: $N = 2 \cdot 11 \cdot 19$, $k = 8$; $\ell = 11$: $N = 2 \cdot 7 \cdot 19$, $k = 12$; $\ell = 19$: $N = 2 \cdot 7 \cdot 11$, $k = 20$; ℓ other: $N = 2 \cdot 7 \cdot 11 \cdot 19$, $k = 2$.

Solution 20. One approach is to view $J_1(N)$ as a complex torus, and note that the endomorphism ring is the set of automorphism of a complex vector space that fix a lattice. Another approach is to use the deeper finiteness theorems that are valid in arbitrary characteristic, see, e.g., [**74**, Thm. 12.5].

CHAPTER 5

Appendix by Brian Conrad: The Shimura construction in weight 2

The purpose of this appendix is to explain the ideas of Eichler-Shimura for constructing the two-dimensional ℓ-adic representations attached to classical weight-2 Hecke eigenforms. We assume familiarity with the theory of schemes and the theory of newforms, but the essential arithmetic ideas are due to Eichler and Shimura. We warn the reader that a complete proof along the lines indicated below requires the verification of a number of compatibilities between algebraic geometry, algebraic topology, and the classical theory of modular forms. As the aim of this appendix is to explain the key arithmetic ideas of the proof, we must pass over in silence the verification of many such compatibilities. However, we at least make explicit what compatibilities we need. To prove them all here would require a serious digression from our expository goal; see [**18**, Ch. 3] for details. It is also worth noting that the form of the arguments we present is *exactly* the weight-2 version of Deligne's more general proof of related results in weight > 1, up to the canonical isomorphism

$$\mathbf{Q}_\ell \otimes_{\mathbf{Z}_\ell} \varprojlim \mathrm{Pic}^0_{X/k}[\ell^n](k) \cong H^1_{\text{ét}}(X, \mathbf{Q}_\ell(1)) \cong H^1_{\text{ét,c}}(Y, \mathbf{Q}_\ell(1))$$

for a proper smooth connected curve X over a separably closed field k of characteristic prime to ℓ, and Y a dense open in X. Using ℓ-adic Tate modules allows us to bypass the general theory of étale cohomology which arises in the case of higher weight.

5.1. Analytic preparations

Fix $i = \sqrt{-1} \in \mathbf{C}$ for all time. Fix an integer $N \geq 5$ and let $X_1(N)^{\text{an}}$ denote the classical analytic modular curve, the "canonical" compactification of $Y_1(N)^{\text{an}} = \Gamma_1(N)\backslash\mathfrak{h}$, where $\mathfrak{h} = \{z \in \mathbf{C} : \mathrm{Im}\, z > 0\}$ and $\Gamma_1(N) \subset \mathrm{SL}_2(\mathbf{Z})$ acts on the left via linear fractional transformations. The classical theory identifies the $\mathbf{C}$-vector space $H^0(X_1(N)^{\text{an}}, \Omega^1_{X_1(N)^{\text{an}}})$ with $S_2(\Gamma_1(N), \mathbf{C})$, the space of weight-2 cusp forms. Note that the classical Riemann surface $X_1(N)^{\text{an}}$ has genus 0 if we consider $N < 5$, while $S_2(\Gamma_1(N), \mathbf{C}) = 0$ if $N < 5$. Thus, assuming $N \geq 5$ is harmless for what we will do.

The Hodge decomposition for the compact Riemann surface $X_1(N)^{\text{an}}$ supplies us with an isomorphism of $\mathbf{C}$-vector spaces

$$\begin{aligned} S_2(\Gamma_1(N),\mathbf{C}) \oplus \overline{S_2(\Gamma_1(N),\mathbf{C})} & \\ \cong & H^0(X_1(N)^{\mathrm{an}}, \Omega^1_{X_1(N)^{\mathrm{an}}}) \oplus H^0(X_1(N)^{\mathrm{an}}, \overline{\Omega}^1_{X_1(N)^{\mathrm{an}}}) \\ \xrightarrow{\sim} & H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{C}}) \\ \cong & H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \otimes_{\mathbf{Z}} \mathbf{C} \end{aligned}$$

(where $\underline{A}$ denotes the constant sheaf attached to an abelian group A). This will be called the (weight-2) *Shimura isomorphism*. We want to define "geometric" operations on $H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ which recover the classical Hecke operators on $S_2(\Gamma_1(N), \mathbf{C})$ via the above isomorphism.

The "geometric" (or rather, cohomological) operations we wish to define can be described in two ways. First, we can use explicit matrices and explicit "upper-half plane" models of modular curves. This has the advantage of being concrete, but it provides little conceptual insight and encourages messy matrix calculations. The other point of view is to identify the classical modular curves as the base of certain universal analytic families of (generalized) elliptic curves with level structure. A proper discussion of this latter point of view would take us too far afield, so we will have to settle for only some brief indications along these two lines (though this is how to best verify compatibility with the algebraic theory via schemes).

Choose a matrix $\gamma_n \in \mathrm{SL}_2(\mathbf{Z})$ with $\gamma_n \equiv \left(\begin{smallmatrix} n^{-1} & * \\ 0 & n \end{smallmatrix}\right) \pmod N$, for $n \in (\mathbf{Z}/N\mathbf{Z})^*$. The action of γ_n on $\mathfrak{h}$ induces an action on $Y_1(N)^{\mathrm{an}}$ and even on $X_1(N)^{\mathrm{an}}$. Associating to each $z \in \mathfrak{h}$ the data of the elliptic curve $\mathbf{C}/[1,z] = \mathbf{C}/(\mathbf{Z} + \mathbf{Z}z)$ and the point $1/N$ of exact order N, we may identify $Y_1(N)^{\mathrm{an}}$ as a *set* with the set of isomorphism classes of pairs (E,P) consisting of an elliptic curve E over $\mathbf{C}$ and a point $P \in E$ of exact order N. The map $Y_1(N)^{\mathrm{an}} \to Y_1(N)^{\mathrm{an}}$ induced by γ_n can then described on the underlying set by $(E,P) \mapsto (E, nP)$, so it is "intrinsic", depending only on $n \in (\mathbf{Z}/N\mathbf{Z})^*$. We denote by $I_n : X_1(N)^{\mathrm{an}} \to X_1(N)^{\mathrm{an}}$ the induced map on $X_1(N)^{\mathrm{an}}$. Once this data (E,P) is formulated in a relative context over an analytic base, we could define the analytic map I_n conceptually, without using the matrix γ_n. We ignore this point here.

The map $z \mapsto \frac{-1}{Nz}$ on $\mathfrak{h}$ induces a map $Y_1(N)^{\mathrm{an}} \to Y_1(N)^{\mathrm{an}}$ which extends to $w_N : X_1(N)^{\mathrm{an}} \to X_1(N)^{\mathrm{an}}$. More conceptually and more generally, if $\zeta \in \boldsymbol{\mu}_N(\mathbf{C})$ is a primitive Nth root of unity, consider the rule w_ζ that sends $(E,P) \in Y_1(N)^{\mathrm{an}}$ to $(E/P, P' \mod P)$, where $P' \in E$ has exact order N and $\langle P, P' \rangle_N = \zeta$, with $\langle\ ,\ \rangle_N$ the Weil pairing on N-torsion points (following the sign conventions of [**62, 77**]; opposite the convention of [**109**]). More specifically, on $\mathbf{C}/[1,z]$ we have $\langle \frac{1}{N}, \frac{z}{N} \rangle_N = e^{2\pi i/N}$. The map w_ζ extends to an analytic map $X_1(N)^{\mathrm{an}} \to X_1(N)^{\mathrm{an}}$. When $\zeta = e^{2\pi i/N}$, we have $w_\zeta = w_N$ due to the above sign convention.

We have induced pullback maps

$$w_\zeta^*, I_n^* : H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \to H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}).$$

We write $\langle n \rangle^*$ rather than I_n^*.

Finally, choose a prime p. Define $\Gamma_1(N,p) \subset \mathrm{SL}_2(\mathbf{Z})$ to be $\Gamma_1(N,p) = \Gamma_1(N) \cap \Gamma_0(p)$ when $p \nmid N$ and $\Gamma_1(N,p) = \Gamma_1(N) \cap \Gamma_0(p)^t$ when $p \mid N$, where the group $\Gamma_0(p)^t$ is the transpose of $\Gamma_0(p)$. Define $Y_1(N,p)^{\mathrm{an}} = \Gamma_1(N,p)\backslash\mathfrak{h}$ and let $X_1(N,p)^{\mathrm{an}}$

be its "canonical" compactification. Using the assignment

$$z \mapsto (\mathbf{C}/[1, z], \frac{1}{N}, \langle\frac{1}{p}\rangle)$$

when $p \nmid N$ and

$$z \mapsto (\mathbf{C}/[1, z], \frac{1}{N}, \langle\frac{z}{p}\rangle)$$

when $p \mid N$, we may identify the *set* $Y_1(N, p)^{\mathrm{an}}$ with the set of isomorphism classes of triples (E, P, C) where $P \in E$ has exact order N and $C \subset E$ is a cyclic subgroup of order p, meeting $\langle P\rangle$ trivially (a constraint if $p \mid N$). Here and below, we denote by $\langle P\rangle$ the (cyclic) subgroup generated by P.

There are unique analytic maps

$$\pi_1^{(p)}, \pi_2^{(p)} : X_1(N, p)^{\mathrm{an}} \to X_1(N)^{\mathrm{an}}$$

determined on $Y_1(N, p)^{\mathrm{an}}$ by

$$\pi_1^{(p)}(E, P, C) = (E, P)$$

and

$$\pi_2^{(p)}(E, P, C) = (E/C, P \bmod C).$$

For example, $\pi_1^{(p)}$ is induced by $z \mapsto z$ on $\mathfrak{h}$, in terms of the above upper half plane uniformization of $Y_1(N)^{\mathrm{an}}$ and $Y_1(N, p)^{\mathrm{an}}$.

We define

$$T_p^* = (\pi_1^{(p)})_* \circ (\pi_2^{(p)})^* : H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \to H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$$

where $(\pi_1^{(p)})_* : H^1(X_1(N, p)^{\mathrm{an}}, \underline{\mathbf{Z}}) \to H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ is the canonical trace map associated to the finite map $\pi_1^{(p)}$ of compact Riemann surfaces. More specifically, we have a canonical isomorphism

$$H^1(X_1(N, p)^{\mathrm{an}}, \underline{\mathbf{Z}}) \cong H^1(X_1(N)^{\mathrm{an}}, (\pi_1^{(p)})_*\underline{\mathbf{Z}})$$

since $(\pi_1^{(p)})_*$ is exact on abelian sheaves, and there is a unique trace map of sheaves $(\pi_1^{(p)})_*\underline{\mathbf{Z}} \to \underline{\mathbf{Z}}$ determined on stalks at $x \in X_1(N)^{\mathrm{an}}$ by

$$\begin{aligned} \prod_{\pi_1^{(p)}(y)=x} \mathbf{Z} &\to \mathbf{Z} \\ (a_y) &\mapsto \Sigma_y e_y a_y \end{aligned} \tag{5.1}$$

where e_y is the ramification degree of y over x via $\pi_1^{(p)}$.

A fundamental compatibility, whose proof we omit for reasons of space, is:

Theorem 5.1. *The weight-2 Shimura isomorphism*

$$\mathrm{Sh}_{\Gamma_1(N)} : S_2(\Gamma_1(N), \mathbf{C}) \oplus \overline{S_2(\Gamma_1(N), \mathbf{C})} \cong H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \otimes_{\mathbf{Z}} \mathbf{C}$$

from (5.1) *identifies* $\langle n\rangle \oplus \overline{\langle n\rangle}$ *with* $\langle n\rangle^* \otimes 1$, $T_p \oplus \overline{T}_p$ *with* $T_p^* \otimes 1$, *and* $w_N \oplus \overline{w}_N$ *with* $w^*_{e^{2\pi i/N}} \otimes 1$.

Let $\mathbf{T}_1(N) \subset \mathrm{End}_{\mathbf{Z}}(H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}))$ be the subring generated by the T_p^*'s and $\langle n\rangle^*$'s. By Theorem 5.1, this is identified via the Shimura isomorphism with the classical (weight-2) Hecke ring at level N. In particular, this ring is commutative (which can be seen directly via cohomological considerations as well). It is clearly a finite flat $\mathbf{Z}$-algebra.

The natural map

$$\mathbf{T}_1(N) \otimes_{\mathbf{Z}} \mathbf{C} \hookrightarrow \mathrm{End}_{\mathbf{C}}(H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \otimes_{\mathbf{Z}} \mathbf{C}) \tag{5.2}$$

induces an *injection* $\mathbf{T}_1(N) \otimes \mathbf{C} \hookrightarrow \mathrm{End}_{\mathbf{C}}(S_2(\Gamma_1(N), \mathbf{C}))$, by Theorem 5.1. This is the classical realization of Hecke operators in weight 2.

Another compatibility we need is between the cup product on $H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ and the (non-normalized) Petersson product on $S_2(\Gamma_1(N), \mathbf{C})$. To be precise, we define an isomorphism $H^2(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \cong \mathbf{Z}$ using the i-orientation of the complex manifold $X_1(N)^{\mathrm{an}}$ (i.e., the "$i\mathrm{d}z \wedge \mathrm{d}\overline{z}$" orientation), so we get via cup product a (perfect) pairing

$$(\ ,\ )_{\Gamma_1(N)} : H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \otimes_{\mathbf{Z}} H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \to H^2(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \cong \mathbf{Z}.$$

This induces an analogous pairing after applying $\otimes_{\mathbf{Z}}\mathbf{C}$. For $f, g \in S_2(\Gamma_1(N), \mathbf{C})$ we define

$$\langle f, g\rangle_{\Gamma_1(N)} = \int_{\Gamma_1(N)\backslash \mathfrak{h}} f(z)\overline{g}(z)\mathrm{d}x\mathrm{d}y$$

where this integral is absolutely convergent since f and g have exponential decay near the cusps. This is a perfect Hermitian pairing.

Theorem 5.2. *Under the weight-2 Shimura isomorphism* $\mathrm{Sh}_{\Gamma_1(N)}$,

$$\big(\mathrm{Sh}_{\Gamma_1(N)}(f_1 + \overline{g}_1), \mathrm{Sh}_{\Gamma_1(N)}(f_2 + \overline{g}_2)\big)_{\Gamma_1(N)} = 4\pi \cdot (\langle f_1, g_2\rangle_{\Gamma_1(N)} - \langle f_2, g_1\rangle_{\Gamma_1(N)}).$$

Note that *both* sides are antilinear in g_1, g_2 and alternating with respect to interchanging the pair (f_1, g_1) and (f_2, g_2). The extra factor of 4π is harmless for our purposes since it does not affect formation of adjoints. What is important is that in the classical theory, conjugation by the involution w_N takes each $T \in \mathbf{T}_1(N)$ to its adjoint with respect to the Petersson product. The most subtle case of this is $T = T_p^*$ for $p \mid N$. For $p \nmid N$ the adjoint of T_p^* is $\langle p^{-1}\rangle^* T_p^*$ and the adjoint of $\langle n\rangle^*$ is $\langle n^{-1}\rangle^*$. These classical facts (especially for T_p^* with $p \mid N$) yield the following important corollary of Theorem 5.2.

Corollary 5.3. *With respect to the pairing* $[x, y]_{\Gamma_1(N)} = (x, w_\zeta^* y)_{\Gamma_1(N)}$ *with* $\zeta = e^{2\pi i/N}$, *the action of* $\mathbf{T}_1(N)$ *on* $H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ *is equivariant. That is,*

$$[x, Ty]_{\Gamma_1(N)} = [Tx, y]_{\Gamma_1(N)}$$

for all $T \in \mathbf{T}_1(N)$. *With respect to* $(\ ,\ )_{\Gamma_1(N)}$, *the adjoint of* T_p^* *for* $p \nmid N$ *is* $\langle p^{-1}\rangle^* T_p^*$ *and the adjoint of* $\langle n\rangle^*$ *is* $\langle n^{-1}\rangle^*$ *for* $n \in (\mathbf{Z}/N\mathbf{Z})^*$.

Looking back at the "conceptual" definition of w_ζ^* for an arbitrary primitive Nth root of unity $\zeta \in \boldsymbol{\mu}_N(\mathbf{C})$, which gives an analytic involution of $X_1(N)^{\mathrm{an}}$, one can check that $w_{\zeta^j}^* \circ w_\zeta^* = \langle j\rangle^*$ for $j \in (\mathbf{Z}/N\mathbf{Z})^*$. Since $\langle j\rangle^*$ is a unit in $\mathbf{T}_1(N)$ and $\mathbf{T}_1(N)$ is *commutative*, we conclude that Corollary 5.3 is true with $\zeta \in \boldsymbol{\mu}_N(\mathbf{C})$ *any* primitive Nth root of unity (by reduction to the case $\zeta = e^{2\pi i/N}$).

Our final step on the analytic side is to reformulate everything above in terms of Jacobians. For any compact Riemann surface X, there is an isomorphism of complex Lie groups

$$\mathrm{Pic}^0_X \cong H^1(X, \mathcal{O}_X)/H^1(X, \underline{\mathbf{Z}}) \tag{5.3}$$

via the exponential sequence

$$0 \to \underline{\mathbf{Z}} \to \mathcal{O}_X \xrightarrow{e^{2\pi i(\cdot)}} \mathcal{O}_X^* \to 1$$

and the identification of the underlying group of Pic^0_X with

$$H^1(X, \mathcal{O}_X^*) \cong \check{H}^1(X, \mathcal{O}_X^*)\,,$$

where the line bundle $\mathcal{L}$ with trivializations $\varphi_i : \mathcal{O}_{U_i} \cong \mathcal{L}|U_i$ corresponds to the class of the Čech 1-cocycle

$$\{\varphi_j^{-1} \circ \varphi_i : \mathcal{O}_{U_i \cap U_j} \cong \mathcal{O}_{U_i \cap U_j}\} \in \prod_{i<j} H^0(U_i \cap U_j, \mathcal{O}_X^*)$$

for an ordered open cover $\{U_i\}$. Beware that the tangent space isomorphism

$$T_0(\mathrm{Pic}^0_X) \cong H^1(X, \mathcal{O}_X)$$

coming from (5.3) is $-2\pi i$ times the "algebraic" isomorphism arising from

$$0 \to \mathcal{O}_X \to \mathcal{O}^*_{X[\varepsilon]} \to \mathcal{O}_X^* \to 1,$$

where $X[\varepsilon] = (X, \mathcal{O}_X[\varepsilon]/\varepsilon^2)$ is the non-reduced space of "dual numbers over X". This extra factor of $-2\pi i$ will not cause problems. We will use (5.3) to "compute" with Jacobians.

Let $f : X \to Y$ be a finite map between compact Riemann surfaces. Since f is finite flat, there is a natural trace map $f_*\mathcal{O}_X \to \mathcal{O}_Y$, and it is not difficult to check that this is compatible with the trace map $f_*\underline{\mathbf{Z}} \to \underline{\mathbf{Z}}$ as defined in (5.1). In particular, we have a trace map

$$f_* : H^1(X, \mathcal{O}_X) \cong H^1(Y, f_*\mathcal{O}_X) \to H^1(Y, \mathcal{O}_Y).$$

Likewise, we have compatible pullback maps $f^*\mathcal{O}_Y \cong \mathcal{O}_X$ and $f^*\underline{\mathbf{Z}} \cong \underline{\mathbf{Z}}$.

Thus, any such f gives rise to *commutative* diagrams

$$\begin{array}{ccc} H^1(Y, \mathcal{O}_Y) & \xrightarrow{f^*} & H^1(X, \mathcal{O}_X) \\ \uparrow & & \uparrow \\ H^1(Y, \underline{\mathbf{Z}}) & \xrightarrow{f^*} & H^1(X, \underline{\mathbf{Z}}) \end{array} \qquad \begin{array}{ccc} H^1(X, \mathcal{O}_X) & \xrightarrow{f_*} & H^1(Y, \mathcal{O}_Y) \\ \uparrow & & \uparrow \\ H^1(X, \underline{\mathbf{Z}}) & \xrightarrow{f_*} & H^1(Y, \underline{\mathbf{Z}}) \end{array},$$

where the columns are induced by the canonical maps $\underline{\mathbf{Z}} \to \mathcal{O}_Y$ and $\underline{\mathbf{Z}} \to \mathcal{O}_X$. Passing to quotients on the columns therefore gives rise to maps

$$f^* : \mathrm{Pic}^0_Y \to \mathrm{Pic}^0_X, \qquad f_* : \mathrm{Pic}^0_X \to \mathrm{Pic}^0_Y$$

of analytic Lie groups. These maps are "computed" by

Lemma 5.4. *In the above situation, $f^* = \mathrm{Pic}^0(f)$ is the map induced by* Pic^0 *functoriality and $f_* = \mathrm{Alb}(f)$ is the map induced by Albanese functoriality. These are dual with respect to the canonical autodualities of* Pic^0_X, Pic^0_Y.

The significance of the theory of Jacobians is that by (5.3) we have a canonical isomorphism

$$\begin{aligned} T_\ell(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}) &\cong H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}_\ell) \\ &\cong H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}}) \otimes_{\mathbf{Z}} \mathbf{Z}_\ell, \end{aligned} \tag{5.4}$$

connecting the ℓ-adic Tate module of $\mathrm{Pic}^0_{X_1(N)}$ with the $\mathbf{Z}$-module $H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ that "encodes" $S_2(\Gamma_1(N), \mathbf{C})$ via the Shimura isomorphism. Note that this isomorphism is defined in terms of the analytic construction (5.3) which depends upon the choice of i. The intrinsic isomorphism (compatible with étale cohomology) has $\mathbf{Z}$ above replaced by $2\pi i\mathbf{Z} = -2\pi i\mathbf{Z}$.

Definition 5.5. We define endomorphisms of $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$ via

$$T_p^* = \mathrm{Alb}(\pi_1^{(p)}) \circ \mathrm{Pic}^0(\pi_2^{(p)}), \quad \langle n\rangle^* = \mathrm{Pic}^0(I_n), \quad w_\zeta^* = \mathrm{Pic}^0(w_\zeta).$$

By Lemma 5.4, it follows that the above isomorphism (5.4) carries the operators on $T_\ell(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}})$ over to the ones *previously defined* on $H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Z}})$ (which are, in turn, compatible with the classical operations via the Shimura isomorphism). By the faithfulness of the "Tate module" functor on complex tori, we conclude that $\mathbf{T}_1(N)$ *acts* on $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$ in a unique manner compatible with the above definition, and (5.4) is an isomorphism of $\mathbf{T}_1(N) \otimes_{\mathbf{Z}} \mathbf{Z}_\ell$-modules. We call this the $(\,)^*$*-action* of $\mathbf{T}_1(N)$ on $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$.

We must warn the reader that under the canonical isomorphism of $\mathbf{C}$-vector spaces

$$\begin{aligned} S_2(\Gamma_1(N), \mathbf{C}) &\cong H^0(X_1(N)^{\mathrm{an}}, \Omega^1_{X_1(N)^{\mathrm{an}}}) \\ &\cong H^0(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}, \Omega^1_{\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}}) \\ &\cong \mathrm{Cot}_0(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}), \end{aligned}$$

the $(\,)^*$-action of $T \in \mathbf{T}_1(N)$ on $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$ does *not* go over to the classical action of T on $S_2(\Gamma_1(N), \mathbf{C})$, but rather the adjoint of T with respect to the Petersson pairing. To clear up this matter, we make the following definition:

Definition 5.6.

$$(T_p)_* = \mathrm{Alb}(\pi_2^{(p)}) \circ \mathrm{Pic}^0(\pi_1^{(p)}), \quad \langle n\rangle_* = \mathrm{Alb}(I_n), \quad (w_\zeta)_* = \mathrm{Alb}(w_\zeta).$$

Since $I_n^{-1} = I_{n^{-1}}$ and $w_\zeta^{-1} = w_\zeta$ on $X_1(N)^{\mathrm{an}}$, we have $(w_\zeta)_* = w_\zeta^*$ and $\langle n\rangle_* = \langle n^{-1}\rangle^*$. We claim that the above $(\,)_*$ operators are the *dual* morphisms (with respect to the canonical principal polarization of $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$) of the $(\,)^*$ operators and induce exactly the *classical* action of T_p and $\langle n\rangle$ on $S_2(\Gamma_1(N), \mathbf{C})$, so we also have a well-defined $(\,)_*$-action of $\mathbf{T}_1(N)$ on $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$, dual to the $(\,)^*$-action. By Theorem 5.2, Corollary 5.3, and Lemma 5.4, this follows from the following general fact about compact Riemann surfaces. The proof is non-trivial.

Lemma 5.7. *Let X be a compact Riemann surface, and use the i-orientation to define $H^2(X, \underline{\mathbf{Z}}) \cong \mathbf{Z}$. Use $1 \mapsto e^{2\pi i/\ell^n}$ to define $\mathbf{Z}/\ell^n \cong \mu_{\ell^n}(\mathbf{C})$ for all n. The*

diagram

$$\begin{array}{ccc} H^1(X,\underline{\mathbf{Z}}_\ell)\otimes_{\mathbf{Z}_\ell} H^1(X,\underline{\mathbf{Z}}_\ell) & \xrightarrow{\cup} & \mathbf{Z}_\ell \\ \downarrow\cong & & \downarrow\cong \\ T_\ell(\mathrm{Pic}^0_X)\otimes_{\mathbf{Z}_\ell} T_\ell(\mathrm{Pic}^0_X) & \longrightarrow & \varprojlim \mu_{\ell^n}(\mathbf{C}) \end{array}$$

anticommutes (i.e., going around from upper left to lower right in the two possible ways gives results that are negatives of each other), where the bottom row is the ℓ-adic Weil pairing (with respect to the canonical principal polarization $\mathrm{Pic}^0_X \cong \widehat{\mathrm{Pic}^0_X}$ for the "second" Pic^0_X in the lower left.)

Note that the sign doesn't affect formation of adjoints. It ultimately comes from the sign on the bottom of [**77**, pg. 237] since our Weil pairing sign convention agrees with [**77**].

We now summarize our findings in terms of $V_\ell(N) = \mathbf{Q}_\ell \otimes_{\mathbf{Z}_\ell} T_\ell(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}})$, which has a perfect alternating Weil pairing

$$(\ ,\ )_\ell : V_\ell(N)\otimes V_\ell(N) \to \mathbf{Q}_\ell(1)$$

and has two $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$-actions, via the $(\ )^*$-actions and the $(\ )_*$-actions. Since $(w_\zeta)_* = w_\zeta^*$, we simply write w_ζ for this operator on $V_\ell(N)$.

Theorem 5.8. *Let $\mathbf{T}_1(N)$ act on $V_\ell(N)$ with respect to the $(\ )^*$-action or with respect to the $(\ )_*$-action. With respect to $(\ ,\ )_\ell$, the adjoint of T_p for $p \nmid N$ is $\langle p\rangle^{-1}T_p$ and the adjoint of $\langle n\rangle$ is $\langle n\rangle^{-1}$ for $n \in (\mathbf{Z}/N\mathbf{Z})^*$. With respect to*

$$[x,y]_\ell = (x, w_\zeta(y))_\ell$$

for $\zeta \in \boldsymbol{\mu}_N(\mathbf{C})$ a primitive Nth root of unity, the action of $\mathbf{T}_1(N)$ on $V_\ell(N)$ is self-adjoint. In general, adjointness with respect to $(\ ,\ )_\ell$ interchanges the $(\ )_*$-action and $(\ )^*$-action.

It should be noted that when making the translation to étale cohomology, the $(\ )^*$-action plays a more prominent role (since this is what makes (5.4) a $\mathbf{T}_1(N)$-equivariant map). However, when working directly with Tate modules and arithmetic Frobenius elements, it is the $(\ )_*$-action which gives the cleaner formulation of Shimura's results.

An important consequence of Theorem 5.8 is

Corollary 5.9. *The $\mathbf{Q}_\ell \otimes_{\mathbf{Z}} \mathbf{T}_1(N)$-module $V_\ell(N)$ is free of rank 2 for either action, and $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q}\otimes\mathbf{T}_1(N),\mathbf{Q})$ is free of rank 1 over $\mathbf{Q}\otimes\mathbf{T}_1(N)$ (hence likewise with $\mathbf{Q}$ replaced by any field of characteristic 0).*

Remark 5.10. The assertion about $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q}\otimes\mathbf{T}_1(N),\mathbf{Q})$ is equivalent to the intrinsic condition that $\mathbf{Q}\otimes\mathbf{T}_1(N)$ is *Gorenstein*. Also, this freeness clearly makes the two assertions about $V_\ell(N)$ for the $(\ )_*$- and $(\ )^*$-actions equivalent. *For the proof*, the $(\ )^*$-action is what we use. But in what follows, it is the case of the $(\ )_*$-action that we need!

Proof. Using (5.4) and the choice of $(\ )^*$-action on $V_\ell(N)$, it suffices to prove

- $H^1(X_1(N)^{\mathrm{an}},\underline{\mathbf{Q}})$ is free of rank 2 over $\mathbf{Q}\otimes\mathbf{T}_1(N)$,
- $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q}\otimes\mathbf{T}_1(N),\mathbf{Q})$ is free of rank 1 over $\mathbf{Q}\otimes\mathbf{T}_1(N)$.

Using $[\,,\,]_{\Gamma_1(N)}$, we have

$$H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Q}}) \cong \mathrm{Hom}_{\mathbf{Q}}(H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Q}}), \mathbf{Q}) \tag{5.5}$$

as $\mathbf{Q} \otimes \mathbf{T}_1(N)$-modules, so we may study this $\mathbf{Q}$-dual instead. Since $\mathbf{Q} \otimes \mathbf{T}_1(N)$ is semilocal, a finite module over this ring is locally free of constant rank if and only if it is *free* of that rank. But local freeness of constant rank can be checked after faithfully flat base change. Applying this with the base change $\mathbf{Q} \to \mathbf{C}$, and noting that $\mathbf{C} \otimes \mathbf{T}_1(N)$ is semilocal, it suffices to replace $\mathbf{Q}$ by $\mathbf{C}$ above.

Note that *if* the right hand side of (5.5) is free of rank 2, so is the left side, so choosing a basis of the left side and feeding it into the right hand side shows that $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q} \otimes \mathbf{T}_1(N)^{\oplus 2}, \mathbf{Q})$ is free of rank 2. In particular, the direct summand $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q} \otimes \mathbf{T}_1(N), \mathbf{Q})$ is flat over $\mathbf{Q} \otimes \mathbf{T}_1(N)$ with full support over $\mathrm{Spec}(\mathbf{Q} \otimes \mathbf{T}_1(N))$, so it must be locally free with local rank at least 1 at all points of $\mathrm{Spec}(\mathbf{Q} \otimes \mathbf{T}_1(N))$. Consideration of $\mathbf{Q}$-dimensions then forces $\mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q} \otimes \mathbf{T}_1(N), \mathbf{Q})$ to be locally free of rank 1, hence free of rank 1. In other words, it suffices to show that $\mathrm{Hom}_{\mathbf{Q}}(H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{Q}}), \mathbf{Q})$ is free of rank 2 over $\mathbf{T}_1(N) \otimes \mathbf{Q}$, or equivalently that $\mathrm{Hom}_{\mathbf{C}}(H^1(X_1(N)^{\mathrm{an}}, \underline{\mathbf{C}}), \mathbf{C})$ is free of rank 2 over $\mathbf{T}_1(N) \otimes \mathbf{C}$.

Via the Shimura isomorphism (in weight 2), which is compatible with the Hecke actions, we are reduced to showing that $\mathrm{Hom}(S_2(\Gamma_1(N), \mathbf{C}), \mathbf{C})$ is free of rank 1 over $\mathbf{C} \otimes \mathbf{T}_1(N)$. For this purpose, we will study the $\mathbf{C} \otimes \mathbf{T}_1(N)$-equivariant $\mathbf{C}$-bilinear pairing

$$\begin{gathered} S_2(\Gamma_1(N), \mathbf{C}) \otimes_{\mathbf{C}} (\mathbf{C} \otimes \mathbf{T}_1(N)) \to \mathbf{C} \\ (f, T) \mapsto a_1(Tf) \end{gathered}$$

were $a_1(\cdot)$ is the "Fourier coefficient of q". This is $\mathbf{C} \otimes \mathbf{T}_1(N)$-equivariant, since $\mathbf{T}_1(N)$ is commutative. It suffices to check that there's no nonzero kernel on either side of this pairing. Since

$$\mathbf{C} \otimes \mathbf{T}_1(N) \to \mathrm{End}_{\mathbf{C}}(S_2(\Gamma_1(N), \mathbf{C}))$$

is *injective* (as noted in (5.2)) and $a_1(TT_nf) = a_n(Tf)$ for $T \in \mathbf{T}_1(N)$, the kernel on the right is trivial. Since $a_1(T_nf) = a_n(f)$, the kernel on the left is also trivial. □

5.2. Algebraic preliminaries

Let S be a scheme. An *elliptic curve* $E \to S$ is a proper smooth group scheme with geometrically connected fibers of dimension 1 (necessarily of genus 1). It follows from [**62**, Ch.2] that the group structure is commutative and uniquely determined by the identity section. Fix $N \geq 1$ and assume $N \in H^0(S, \mathcal{O}_S^*)$ (i.e., S is a $\mathbf{Z}[\frac{1}{N}]$-scheme). Thus, the map $N : E \to E$ is finite *étale* of degree N^2 as can be checked on geometric fibers. A *point of exact order* N on E is a section $P : S \to E$ which is killed by N (i.e., factors through the finite étale group scheme $E[N]$) and induces a point of exact order N on geometric fibers.

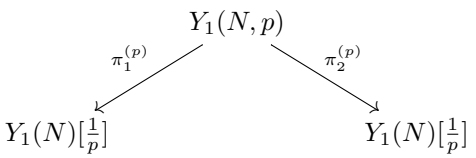
It follows from the stack-theoretic methods in [**25**] or the more explicit descent arguments in [**62**] that for $N \geq 5$ there is a proper smooth $\mathbf{Z}[\frac{1}{N}]$-scheme $X_1(N)$ equipped with a finite flat map to $\mathbf{P}^1_{\mathbf{Z}[\frac{1}{N}]}$, such that the open subscheme $Y_1(N)$ lying over $\mathbf{P}^1_{\mathbf{Z}[\frac{1}{N}]} - \{\infty\} = \mathbf{A}^1_{\mathbf{Z}[\frac{1}{N}]}$ is the base of a universal object $(E_1(N), P) \to Y_1(N)$ for elliptic curves with a point of exact order N over variable $\mathbf{Z}[\frac{1}{N}]$-schemes.

Moreover, the fibers of $X_1(N) \to \mathrm{Spec}\, \mathbf{Z}[\frac{1}{N}]$ are *geometrically connected*, as this can be checked on a single geometric fiber and by choosing the complex fiber we may

appeal to the fact (whose proof requires some care) that there is an isomorphism $(X_1(N) \times_{\mathbf{Z}[\frac{1}{N}]} \mathbf{C})^{\mathrm{an}} \cong X_1(N)^{\mathrm{an}}$ identifying the "algebraic" data $(\mathbf{C}/[1,z],\frac{1}{N})$ in $Y_1(N)(\mathbf{C}) \subset X_1(N)(\mathbf{C})$ with the class of $z \in \mathfrak{h}$ in $\Gamma_1(N)\backslash\mathfrak{h} = Y_1(N)^{\mathrm{an}} \subset X_1(N)^{\mathrm{an}}$ (and $X_1(N)^{\mathrm{an}}$ *is* connected, as $\mathfrak{h}$ is). These kinds of compatibilities are somewhat painful to check unless one develops a full-blown relative theory of elliptic curves in the analytic world (in which case the verifications become quite mechanical and natural).

Again fixing $N \geq 5$, but now also a prime p, we want an algebraic analogue of $X_1(N,p)^{\mathrm{an}}$ over $\mathbf{Z}[\frac{1}{Np}]$. Let $(E,P) \to S$ be an elliptic curve with a point of exact order N over a $\mathbf{Z}[\frac{1}{Np}]$-scheme S. We're interested in studying triples $(E,P,C) \to S$ where $C \subset E$ is an order-p finite locally free S-subgroup-scheme which is not contained in the subgroup generated by P on geometric fibers (if $p \mid N$). Methods in [**25**] and [**62**] ensure the existence of a universal such object $(E_1(N,p),P,C) \to Y_1(N,p)$ for a smooth affine $\mathbf{Z}[\frac{1}{Np}]$-scheme which naturally sits as the complement of a relative Cartier divisor in a proper smooth $\mathbf{Z}[\frac{1}{Np}]$-scheme $X_1(N,p)$ which is finite flat over $\mathbf{P}^1_{\mathbf{Z}[\frac{1}{Np}]}$ (with $Y_1(N,p)$ the preimage of $\mathbf{A}^1_{\mathbf{Z}[\frac{1}{Np}]}$). Base change to $\mathbf{C}$ and analytification recovers $X_1(N,p)^{\mathrm{an}}$ as before, so $X_1(N,p) \to \operatorname{Spec} \mathbf{Z}[\frac{1}{Np}]$ has geometrically connected fibers.

There are maps of $\mathbf{Z}[\frac{1}{Np}]$-schemes (respectively, $\mathbf{Z}[\frac{1}{N}]$-schemes)



$$Y_1(N) \xrightarrow{I_n} Y_1(N)$$

determined by $(E,P,C) \xrightarrow{\pi_1^{(p)}} (E,P)$ and $(E,P,C) \xrightarrow{\pi_2^{(p)}} (E/C,P)$ (which makes sense in $Y_1(N)$ if $p \mid N$ by the "disjointness" condition on C and P) and $I_n(E,P) = (E,nP)$. Although $\pi_2^{(p)}$ is *not* a map over $\mathbf{A}^1_{\mathbf{Z}[\frac{1}{Np}]}$, it can be shown that these all uniquely extend to (necessarily finite *flat*) maps, again denoted $\pi_1^{(p)}$, $\pi_2^{(p)}$, I_n between $X_1(N,p)$, $X_1(N)[\frac{1}{p}]$, $X_1(N)$. A proof of this fact requires the theory of minimal regular proper models of curves over a Dedekind base; the analogous fact over $\mathbf{Q}$ is an immediate consequence of basic facts about proper smooth curves over a field, but in order to most easily do some later calculations in characteristic $p \nmid N$ it is convenient to know that we have the map I_p defined on $X_1(N)$ over $\mathbf{Z}[1/N]$ (though this could be bypassed by using liftings to characteristic 0 in a manner similar to our later calculations of T_p in characteristic p).

Likewise, over $\mathbf{Z}[\frac{1}{N},\zeta_N]$ we can define, for any primitive Nth root of unity $\zeta = \zeta_N^i$ ($i \in (\mathbf{Z}/N\mathbf{Z})^*$), an operator $w_\zeta : Y_1(N)_{/\mathbf{Z}[\frac{1}{N},\zeta_N]} \to Y_1(N)_{/\mathbf{Z}[\frac{1}{N},\zeta_N]}$ via $w_\zeta(E,P) = (E/\langle P\rangle, P')$ where $\langle P\rangle$ is the order-N étale subgroup-scheme generated by P and $P' \in (E[N]/\langle P\rangle)(S)$ is uniquely determined by the relative Weil pairing condition $\langle P,P'\rangle_N = \zeta$ (with $P' \in E[N](S)$ here). This really does extend to $X_1(N)_{/\mathbf{Z}[\frac{1}{N},\zeta_N]}$, and one checks that $w_{\zeta^j} w_\zeta = I_j$ for $j \in (\mathbf{Z}/N\mathbf{Z})^*$. In particular, $w_\zeta^2 = 1$.

Since $X_1(N) \to \operatorname{Spec} \mathbf{Z}[\frac{1}{N}]$ is a proper smooth scheme with geometrically connected fibers of dimension 1, $\operatorname{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$ is an abelian scheme over $\mathbf{Z}[\frac{1}{N}]$ and

hence is the Néron model of its generic fiber. We have scheme-theoretic Albanese and Pic^0 functoriality for finite (flat) maps between proper smooth curves (with geometrically connected fibers) over any base at all, and analytification of such a situation over $\mathbf{C}$ recovers the classical theory of Pic^0 as used in Section 5.1.

For example, we have endomorphisms

$$\langle n\rangle^* = \mathrm{Pic}^0(I_n), \quad \langle n\rangle_* = \mathrm{Alb}(I_n)$$

on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$,

$$w_\zeta^* = \mathrm{Pic}^0(w_\zeta) = \mathrm{Alb}(w_\zeta) = (w_\zeta)_*$$

on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N},\zeta_N]}}$, and

$$\begin{aligned} T_p^* &= \mathrm{Alb}(\pi_1^{(p)}) \circ \mathrm{Pic}^0(\pi_2^{(p)}) \\ (T_p)_* &= \mathrm{Alb}(\pi_2^{(p)}) \circ \mathrm{Pic}^0(\pi_1^{(p)}) \end{aligned}$$

on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{Np}]}}$. A key point is that by the *Néronian property*, T_p^* and $(T_p)_*$ uniquely extend to endomorphisms of $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$, even though the $\pi_i^{(p)}$ do *not* make sense over $\mathbf{Z}[\frac{1}{N}]$ from what has gone before. In particular, it makes sense to study T_p^* and $(T_p)_*$ on the abelian variety $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{F}_p}}$ over $\mathbf{F}_p$ for $p \nmid N$. This will be rather crucial later, but note it requires the Néronian property in the definition.

Passing to the analytifications, the above constructions recover the operators defined on $\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}$ in Section 5.1. The resulting subring of

$$\mathrm{End}(\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}) \subset \mathrm{End}(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}})$$

generated by T_p^*, $\langle n\rangle^*$ (respectively, by $(T_p)_*$, $\langle n\rangle_*$) is identified with $\mathbf{T}_1(N)$ via its $(\,)^*$-action (respectively, via its $(\,)_*$-action) and using

$$\varprojlim \mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}[\ell^n](\overline{\mathbf{Q}}) \cong T_\ell(\mathrm{Pic}^0_{X_1(N)^{\mathrm{an}}}) \tag{5.6}$$

(using $\overline{\mathbf{Q}} \subset \mathbf{C}$) endows our "analytic" $V_\ell(N)$ with a canonical *continuous* action of $G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ unramified at all $p \nmid N\ell$ (via Néron-Ogg-Shafarevich) and *commuting* with the action of $\mathbf{T}_1(N)$ (via either the $(\,)^*$-action or the $(\,)_*$-action). We also have an endomorphism $w_\zeta = w_\zeta^* = (w_\zeta)_*$ on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N},\zeta_N]}}$ and it is easy to see that

$$(g^{-1})^* w_{g(\zeta)} g^* = w_\zeta$$

on $\overline{\mathbf{Q}}$-points, where $g \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and g^* denotes the natural action of g on $\overline{\mathbf{Q}}$-points (corresponding to base change of degree 0 line bundles on $X_1(N)_{/\overline{\mathbf{Q}}}$). Since $w_\zeta = w_{\zeta^{-1}}$ (as $(E,P) \cong (E,-P)$ via -1), we see that w_ζ is defined over the real subfield $\mathbf{Q}(\zeta_N)^+$. By étale descent, the operator w_ζ is defined over $\mathbf{Z}[\frac{1}{N},\zeta_N]^+$.

In any case, w_ζ acts on $V_\ell(N)$, recovering the operator in Section 5.1, and so this conjugates the $(\,)^*$-action to the $(\,)_*$-action, taking each $T \in \mathbf{T}_1(N)$ (for either action on $V_\ell(N)$) to its Weil pairing adjoint, via the canonical principal polarization of the abelian scheme $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$. Using Corollary 5.3 and (5.6) we obtain

Lemma 5.11. *Let $\mathbf{T}_1(N)$ act on $V_\ell(N)$ through either the $(\,)^*$-action or the $(\,)_*$-action. Then $\rho_{N,\ell} : G_{\mathbf{Q}} \to \mathrm{Aut}(V_\ell(N)) \cong \mathrm{GL}(2, \mathbf{Q}_\ell \otimes \mathbf{T}_1(N))$ is a continuous representation, unramified at $p \nmid N\ell$.*

The main result we are after is

Theorem 5.12. *Let $\mathbf{T}_1(N)$ act on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$ via the $(\,)_*$-action. For any $p \nmid N\ell$, the characteristic polynomial of $\rho_{N,\ell}(\mathrm{Frob}_p)$ is*

$$X^2 - (T_p)_* X + p\langle p\rangle_*$$

relative to the $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$-module structure on $V_\ell(N)$, where Frob_p denotes an arithmetic Frobenius element at p.

The proof of Theorem 5.12 will make essential use of the w_ζ operator. For the remainder of this section, we admit Theorem 5.12 and deduce its consequences. Let $f \in S_2(\Gamma_1(N), \mathbf{C})$ be a *newform* of level N. Let $K_f \subset \mathbf{C}$ be the number field generated by $a_p(f)$ for all $p \nmid N$, where $f = \sum a_n(f) q^n$, so by weak multiplicity one $a_n(f) \in K_f$ for all $n \geq 1$ and the Nebentypus character χ_f has values in K_f. Let $\mathfrak{p}_f \subset \mathbf{T}_1(N)$ be the minimal prime corresponding to f (i.e., the kernel of the map $\mathbf{T}_1(N) \to K_f$ sending each $T \in \mathbf{T}_1(N)$ to its eigenvalue on f).

We now require $\mathbf{T}_1(N)$ to act on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}$ via its $(\,)_*$-action.

Definition 5.13. A_f is the quotient of $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Q}}}$ by $\mathfrak{p}_f \subset \mathbf{T}_1(N)$.

By construction, A_f has good reduction over $\mathbf{Z}[\frac{1}{N}]$ and the action of $\mathbf{T}_1(N)$ on $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Q}}}$ induces an action of $\mathbf{T}_1(N)/\mathfrak{p}$ on A_f, hence an action of $K_f \cong (\mathbf{T}_1(N)/\mathfrak{p}) \otimes_{\mathbf{Z}} \mathbf{Q}$ on A_f in the "up-to-isogeny" category.

Theorem 5.14 (Shimura)**.** *We have $\dim A_f = [K_f : \mathbf{Q}]$ and $V_\ell(A_f)$ is free of rank 2 over $\mathbf{Q}_\ell \otimes_{\mathbf{Q}} K_f$, with Frob_p having characteristic polynomial*

$$X^2 - (1 \otimes a_p(f))X + 1 \otimes p\chi_f(p)$$

for all $p \nmid N\ell$.

Proof. By Lemma 5.11 and Theorem 5.12, we just have to check that the $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$-linear map

$$V_\ell(\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Q}}}) \to V_\ell(A_f)$$

identifies the right hand side with the quotient of the left hand side by $\mathfrak{p}_f$. More generally, for any exact sequence

$$B' \to B \to A \to 0$$

of abelian varieties over a field of characteristic prime to ℓ, we claim

$$V_\ell(B') \to V_\ell(B) \to V_\ell(A) \to 0$$

is exact. We may assume the base field is algebraically closed, and then may appeal to Poincaré reducibility (see [**77**, pg. 173]). □

Choosing a place λ of K_f over ℓ and using the natural realization of $K_{f,\lambda}$ as a factor of $\mathbf{Q}_\ell \otimes K_f$, we deduce from Theorem 5.14:

Corollary 5.15. *Let $f \in S_2(\Gamma_1(N), \mathbf{C})$ be a newform and λ a place of K_f over ℓ. There exists a continuous representation*

$$\rho_{f,\lambda} : G_{\mathbf{Q}} \to GL(2, K_{f,\lambda})$$

unramified at all $p \nmid N\ell$, with Frob_p *having characteristic polynomial*

$$X^2 - a_p(f)X + p\chi_f(p) \in K_{f,\lambda}[X].$$

5.3. Proof of Theorem 5.12

Fix $p \nmid N$ and let

$$J_p = \mathrm{Pic}^0_{X_1(N)_{/\mathbf{F}_p}} \cong \mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}} \times_{\mathbf{Z}[\frac{1}{N}]} \mathbf{F}_p$$

with $\mathbf{T}_1(N)$ acting through the $(\,)_*$-action. Fix a choice of Frob_p, or more specifically fix a choice of place in $\overline{\mathbf{Q}}$ over p. Note that this determines a preferred algebraic closure $\overline{\mathbf{F}}_p$ as a quotient of the ring of algebraic integers, and in particular a map $\mathbf{Z}[1/N, \zeta_N] \to \overline{\mathbf{F}}_p$. Thus, we may view w_ζ as inducing an endomorphism of the abelian variety $J_p \times_{\mathbf{F}_p} \overline{\mathbf{F}}_p$ over $\overline{\mathbf{F}}_p$ (whereas the elements in $\mathbf{T}_1(N)$ induce endomorphisms of J_p over $\mathbf{F}_p$). The canonical isomorphism

$$V_\ell(\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Q}}}) \cong V_\ell(\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}) \cong V_\ell(J_p)$$

identifies the Frob_p-action on $\overline{\mathbf{Q}}$-points on the left hand side with the (arithmetic) Frobenius action on $\overline{\mathbf{F}}_p$-points on the right hand side. Obviously $V_\ell(J_p)$ is a module over the ring $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$ and is free of rank 2 as such. For *any* $\mathbf{F}_p$-schemes Z, Z' and any $\mathbf{F}_p$-map $f : Z \to Z'$ the diagram

$$\begin{array}{ccc} Z & \xrightarrow{f} & Z' \\ {\scriptstyle F_Z}\downarrow & & \downarrow{\scriptstyle F_{Z'}} \\ Z & \xrightarrow{f} & Z' \end{array} \tag{5.7}$$

commutes, where columns are absolute Frobenius. Taking $Z = \mathrm{Spec}\,\overline{\mathbf{F}}_p$, $Z' = J_p$, we see that the Frob_p action of $V_\ell(J_p)$ through $\overline{\mathbf{F}}_p$-points is *identical* to the action induced by the intrinsic absolute Frobenius morphism $F : J_p \to J_p$ over $\mathbf{F}_p$. Here is the essential input, to be proven later.

Theorem 5.16 (Eichler-Shimura)**.** *In* $\mathrm{End}_{\overline{\mathbf{F}}_p}(J_p)$,

$$(T_p)_* = F + \langle p \rangle_* F^\vee, \qquad w_\zeta^{-1} F w_\zeta = \langle p \rangle_*^{-1} F$$

where $F^\vee$ denotes the dual morphism.

The extra relation involving w_ζ is crucial. The interested reader should compare this with [**108**, Cor. 7.10].

Let us admit Theorem 5.16 and use it to prove Theorem 5.12. We will then prove Theorem 5.16. Using an $\mathbf{F}_p$-rational base point P (e.g., the cusp 0), we get a commutative diagram

$$\begin{array}{ccc} X_1(N)_{/\mathbf{F}_p} & \hookrightarrow & J_p \\ {\scriptstyle F_{X_1(N)}}\downarrow & & \downarrow{\scriptstyle F} \\ X_1(N)_{/\mathbf{F}_p} & \hookrightarrow & J_p \end{array}$$

where $F_{X_1(N)}$ denotes the absolute Frobenius morphism of $X_1(N)_{/\mathbf{F}_p}$, so by Albanese functoriality $F = \mathrm{Alb}(F_{X_1(N)})$. Thus

$$\begin{aligned} FF^\vee &= \mathrm{Alb}(F_{X_1(N)}) \circ \mathrm{Pic}^0(F_{X_1(N)}) \\ &= \deg(F_{X_1(N)}) = p \end{aligned}$$

as $X_1(N)_{/\mathbf{F}_p}$ is a smooth *curve*. We conclude from $(T_p)_* = F + \langle p\rangle_* F^\vee$ that

$$F^2 - (T_p)_* F + p\langle p\rangle_* = 0$$

on J_p, hence in $V_\ell(J_p)$. Thus, $\rho_{N,\ell}(\mathrm{Frob}_p)$ satisfies the expected quadratic polynomial

$$X^2 - (T_p)_* X + p\langle p\rangle_* = 0.$$

Let $X^2 - aX + b$ be the *true* characteristic polynomial, which $\rho_{N,\ell}(\mathrm{Frob}_p)$ must also satisfy, by Cayley-Hamilton. We must *prove* that $a = (T_p)_*$, and then $b = p\langle p\rangle_*$ is forced. It is this matter which requires the second relation.

We want $\mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(\rho_{N,\ell}(\mathrm{Frob}_p)) = (T_p)_*$ or equivalently

$$\mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(V_\ell(F)) = (T_p)_*.$$

Using the modified Weil pairing

$$[x, y]_\ell = (x, w_\zeta y)_\ell$$

and using the fact that $V_\ell(J_p) \cong V_\ell(\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Q}}})$ respects Weil pairings (by invoking the relativization of this concept, here over $\mathbf{Z}[\frac{1}{N}]$) we may identify (via Theorem 5.8 and a choice $\mathbf{Q}_\ell(1) \cong \mathbf{Q}_\ell$ as $\mathbf{Q}_\ell$-vector spaces)

$$V_\ell(J_p) \cong \mathrm{Hom}_{\mathbf{Q}_\ell}(V_\ell(J_p), \mathbf{Q}_\ell) := V_\ell(J_p)^*$$

as $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$-modules, but taking the F-action over to the $\langle p\rangle_* F^\vee$-action, since adjoints with respect to Weil pairings are dual morphisms and $w_\zeta^{-1} F^\vee w_\zeta$ is dual to $w_\zeta^{-1} F w_\zeta = \langle p\rangle_*^{-1} F = F\langle p\rangle_*^{-1}$ (absolute Frobenius commutes with all morphisms of $\mathbf{F}_p$-schemes!)

Since $V_\ell(J_p)$ is free of rank 2 over $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$ and $\mathrm{Hom}_{\mathbf{Q}_\ell}(\mathbf{Q}_\ell \otimes \mathbf{T}_1(N), \mathbf{Q}_\ell)$ is free of rank 1 over $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$, by Corollary 5.9, we conclude

$$\mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(F | V_\ell(J_p)) = \mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(\langle p\rangle_* F^\vee | V_\ell(J_p)^*).$$

We wish to invoke the following applied to the $\mathbf{Q}_\ell$-algebra $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$ and the $\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)$-module $V_\ell(J_p)$:

Lemma 5.17. *Let $\mathcal{O}$ be a commutative ring, A a finite locally free $\mathcal{O}$-algebra with $\mathrm{Hom}_{\mathcal{O}}(A, \mathcal{O})$ a locally free A-module (necessarily of rank 1). Let M be a finite locally free A-module, $M^* = \mathrm{Hom}_{\mathcal{O}}(M, \mathcal{O})$, so M^* is finite and locally free over A with the same rank as M. For any A-linear map $f : M \to M$ with $\mathcal{O}$-dual $f^* : M^* \to M^*$, automatically A-linear,*

$$\mathrm{char}(f) = \mathrm{char}(f^*)$$

in $A[T]$ (these are the characteristic polynomials).

Proof. Without loss of generality $\mathcal{O}$ is local, so A is semilocal. Making faithfully flat base change to the henselization of $\mathcal{O}$ (or the completion if $\mathcal{O}$ is noetherian or if we first reduce to the noetherian case), we may assume that A is a product of local rings. Without loss of generality, A is then local, so

$$M = \oplus A e_i$$

if free, and $\mathrm{Hom}_{\mathcal{O}}(A,\mathcal{O})$ is free of rank 1 over A. Choose an isomorphism

$$h : A \cong \mathrm{Hom}_{\mathcal{O}}(A,\mathcal{O})$$

as A-modules, so the projections

$$\pi_i : M \to Ae_i \cong A$$

satisfy $e_i^* = h(i) \circ \pi_i$ in M^*. These e_i^* are an A-basis of M^* and we compute matrices over A:

$$\mathrm{Mat}_{\{e_i\}}(f) = \mathrm{Mat}_{\{e_i^*\}}(f^*)^t.$$

□

We conclude that

$$\mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(F|V_\ell(J_p)) = \mathrm{tr}_{\mathbf{Q}_\ell \otimes \mathbf{T}_1(N)}(\langle p\rangle_* f^\vee | V_\ell(J_p)).$$

By Theorem 5.16, we have

$$\begin{aligned} 2(T_p)_* &= \mathrm{tr}((T_p)_*|V_\ell(J_p)) \\ &= \mathrm{tr}(F + \langle p\rangle_* F^\vee | V_\ell(J_p)) \\ &= 2\,\mathrm{tr}(F|V_\ell(J_p)). \end{aligned}$$

This proves that $\mathrm{tr}(F|V_\ell(J_p)) = (T_p)_*$, so indeed $X^2 - (T_p)_* X + p\langle p\rangle_*$ *is* the characteristic polynomial. Finally, there remains

Proof of Theorem 5.16. It suffices to check the maps coincide on a Zariski dense subset of $J_p(\overline{\mathbf{F}}_p) = \mathrm{Pic}^0(X_1(N)_{/\overline{\mathbf{F}}_p})$. If g is the genus of $X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}$ and we fix an $\overline{\mathbf{F}}_p$-rational base point, we get an induced surjective map

$$X_1(N)^g_{/\overline{\mathbf{F}}_p} \to J_{p/\overline{\mathbf{F}}_p},$$

so for any dense open $U \subset X_1(N)_{\overline{\mathbf{F}}_p}$, $U^g \to (J_p)_{/\overline{\mathbf{F}}_p}$ hits a Zariski dense subset of $\overline{\mathbf{F}}_p$-points. Taking U to be the ordinary locus of $Y_1(N)_{/\overline{\mathbf{F}}_p}$, it suffices to study what happens to a difference $(x) - (x')$ for $x, x' \in Y_1(N)(\overline{\mathbf{F}}_p)$ corresponding to (E,P), (E',P') over $\overline{\mathbf{F}}_p$ with E and E' *ordinary* elliptic curves.

By the commutative diagram (5.7), the map

$$J_p(\overline{\mathbf{F}}_p) \to J_p(\overline{\mathbf{F}}_p)$$

induced by F is the same as the map induced by the pth power map in $\overline{\mathbf{F}}_p$. By *definition* of Pic^0 functoriality, this corresponds to base change of an invertible sheaf on $X_1(N)_{/\overline{\mathbf{F}}_p}$ by the absolute Frobenius on $\overline{\mathbf{F}}_p$. By *definition* of $Y_1(N)_{/\overline{\mathbf{F}}_p}$ as a universal object, such base change induces on $Y_1(N)(\overline{\mathbf{F}}_p)$ *exactly* "base change by absolute Frobenius" on elliptic curves with a point of exact order N over $\overline{\mathbf{F}}_p$. We conclude

$$F((x) - (x')) = (E^{(p)}, P^{(p)}) - ((E')^{(p)}, P^{(p)})$$

where $(\,)^{(p)}$ denotes base change by absolute Frobenius on $\overline{\mathbf{F}}_p$.

Since $p = FF^\vee = F^\vee F$ and F is bijective on $\overline{\mathbf{F}}_p$-points, we have

$$\begin{aligned} F^\vee((x) - (x')) &= pF^{-1}((x) - (x')) \\ &= p((E^{(p^{-1})}, P^{(p^{-1})}) - ((E')^{(p^{-1})}, (P')^{(p^{-1})})). \end{aligned}$$

Thus,

$$\langle p\rangle_* F^\vee((x) - (x')) = p(E^{(p^{-1})}, pP^{(p^{-1})}) - p((E')^{(p^{-1})}, p(P')^{(p^{-1})})$$

so

$$(F + \langle p\rangle_* F^\vee)((x) - (x')) = (E^{(p)}, P^{(p)}) + p(E^{(p^{-1})}, pP^{(p^{-1})}) \\ - ((E')^{(p)}, (P')^{(p)}) + p((E')^{(p^{-1})}, p(P')^{(p^{-1})}).$$

Computing $(T_p)_*$ on $J_p = \mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}} \times_{\mathbf{Z}[\frac{1}{N}]} \mathbf{F}_p$ is more subtle because $(T_p)_*$ was defined over $\mathbf{Z}[\frac{1}{Np}]$ (or over $\mathbf{Q}$) as $(\pi_2)_*\pi_1^*$ and was *extended* over $\mathbf{Z}[\frac{1}{N}]$ by the Néronian property. That is, we do *not* have a direct definition of $(T_p)_*$ in characteristic p, so we will need to lift to characteristic 0 to compute. It is *here* that the ordinariness assumption is crucial, for we shall see that, in some sense,

$$(T_p)_*((x) - (x')) = (F + \langle p\rangle_* F^\vee)((x) - (x'))$$

as *divisors* for ordinary points x, x'. This is, of course, much stronger than the mere linear equivalence that we need to prove.

Before we dive into the somewhat subtle calculation of $(T_p)_*((x) - (x'))$, let's quickly take care of the relation $w_\zeta^{-1} F w_\zeta = \langle p\rangle_*^{-1} F$, or equivalently,

$$F w_\zeta = w_\zeta \langle p^{-1}\rangle_* F.$$

All maps here are induced by maps on $X_1(N)_{/\overline{\mathbf{F}}_p}$, with $F = \mathrm{Alb}(F_{X_1(N)})$, $w_\zeta = \mathrm{Alb}(w_{\zeta|_{X_1(N)}})$, $\langle p^{-1}\rangle_* = \mathrm{Alb}(I_{p^{-1}})$. Thus, it suffices to show

$$F_{X_1(N)} \circ w_\zeta = w_\zeta I_{p^{-1}} F_{X_1(N)}$$

on $X_1(N)_{/\overline{\mathbf{F}}_p}$, and we can check by studying $x = (E, P) \in Y_1(N)(\overline{\mathbf{F}}_p)$:

$$F_{X_1(N)} w_\zeta(x) = F_{X_1(N)}(E/P, P') = (E^{(p)}/P^{(p)}, (P')^{(p)})$$

where $\langle P, P'\rangle_N = \zeta$, so $\langle P^{(p)}, (P')^{(p)}\rangle_N = \zeta^p$ by compatibility of the (relative) Weil pairing with respect to base change. Meanwhile,

$$w_\zeta I_{p^{-1}} F_{X_1(N)}(x) = w_\zeta(E^{(p)}, p^{-1}P^{(p)}) = (E^{(p)}/(p^{-1}P^{(p)}), Q)$$

where $\langle p^{-1}P^{(p)}, Q\rangle_N = \zeta$, or equivalently $\langle P^{(p)}, Q\rangle = \zeta^p$. Since $Q = (P')^{(p)}$ is such a point, this second relation is established.

Now we turn to the problem of computing

$$(T_p)_*((x) - (x'))$$

for "ordinary points" $x = (E, P)$, $x' = (E', P')$ as above. Let $R = \mathbf{Z}_p^{\mathrm{un}}$, $W(\overline{\mathbf{F}}_p)$, or more generally any henselian (e.g., complete) discrete valuation ring with residue field $\overline{\mathbf{F}}_p$ and fraction field K of characteristic 0. Since $p \nmid N$, R is a $\mathbf{Z}[\frac{1}{N}]$-algebra. Since $Y_1(N)$ is *smooth* over $\mathbf{Z}[\frac{1}{N}]$, we conclude from the (strict) henselian property that $Y_1(N)(R) \to Y_1(N)(\overline{\mathbf{F}}_p)$ is surjective. Of course, this can be seen "by hand": if (E, P) is given over $\overline{\mathbf{F}}_p$, choose a Weierstrass model $\mathcal{E} \hookrightarrow \mathbf{P}^2_R$ lifting E (this is canonically an elliptic curve, by [**62**, Ch 2]). The finite *étale* group scheme $\mathcal{E}[N]$ is *constant* since R is strictly henselian. Thus there exists a unique closed immersion of group schemes $\mathbf{Z}/N\mathbf{Z} \hookrightarrow \mathcal{E}[N]$ lifting $P : \mathbf{Z}/N\mathbf{Z} \hookrightarrow E[N]$.

Let $(\mathcal{E}, \mathcal{P})$, $(\mathcal{E}', \mathcal{P}')$ over R lift x, x' respectively. We view these sections to $X_1(N)_{/R} \to \mathrm{Spec}\, R$ as relative effective Cartier divisors of degree 1. Using the reduction map

$$\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}(R) \to J_p(\overline{\mathbf{F}}_p)$$

and the *definition* of $(T_p)_*$, we see that $(T_p)_*((x)-(x'))$ is the image of $(T_p)_*((\mathcal{E},\mathcal{P})-(\mathcal{E}',\mathcal{P}'))$. Now R is *NOT* a $\mathbf{Z}[\frac{1}{Np}]$-algebra but K *is*, and we have an injection (even bijection)

$$\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}(R) \hookrightarrow \mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}(K),$$

as $\mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}} \to \mathrm{Spec}\,\mathbf{Z}[\frac{1}{N}]$ is separated (even proper).

Thus, we will first compute $(T_p)_*((x)-(x'))$ by working with $\overline{K}$-points, where $\overline{K}$ is an algebraic closure of K. Since $p \nmid N$, we have

$$(\pi_2)_*\pi_1^*((\mathcal{E},\mathcal{P})_{/\overline{K}}) = \sum_C (\mathcal{E}_{\overline{K}}/C, \mathcal{P}_{\overline{K}} \bmod C)$$

where C runs through all $p+1$ order-p subgroups of $\mathcal{E}_{/\overline{K}}$. Since $\mathcal{E} \to \mathrm{Spec}\,R$ has *ordinary* reduction, and R is strictly henselian, the connected-étale sequence of $\mathcal{E}[p]$ is the short exact sequence of finite flat R-group schemes

$$0 \to \mu_p \to \mathcal{E}[p] \to \underline{\mathbf{Z}/p\mathbf{Z}} \to 0.$$

Enlarging R to a finite extension does not change the residue field $\overline{\mathbf{F}}_p$, so we may assume that

$$\mathcal{E}[p]_{/K} \cong \underline{\mathbf{Z}/p\mathbf{Z}} \times \underline{\mathbf{Z}/p\mathbf{Z}}.$$

Taking the scheme-theoretic closure in $\mathcal{E}[p]$ of the $p+1$ distinct subgroups of $\mathcal{E}[p]_{/K}$ gives $p+1$ *distinct* finite flat subgroup schemes $\mathcal{C} \subset \mathcal{E}$ realizing the $p+1$ distinct C's over $\overline{K}$.

Exactly one of these $\mathcal{C}$'s is killed by $\mathcal{E}[p] \to \underline{\mathbf{Z}/p\mathbf{Z}}$ over R, as this can be checked on the generic fiber, so it must be $\mu_p \hookrightarrow \mathcal{E}[p]$. For the remaining $\mathcal{C}$'s, the map $\mathcal{C} \to \underline{\mathbf{Z}/p\mathbf{Z}}$ is an isomorphism on the generic fiber. We claim these maps

$$\mathcal{C} \to \underline{\mathbf{Z}/p\mathbf{Z}}$$

over R are isomorphisms. Indeed, if $\mathcal{C}$ is *étale* this is clear, yet $\mathcal{C} \hookrightarrow \mathcal{E}[p]$ is a finite flat closed subgroup-scheme of order p, so a consideration of the closed fiber shows that if $\mathcal{C}$ is *not* étale then it is multiplicative. But $\mathcal{E}[p]$ has a *unique* multiplicative subgroup-scheme since

$$\mathcal{E}[p]^\vee \cong \mathcal{E}[p]$$

by Cartier-Nishi duality and $\mathcal{E}[p]$ has a *unique* order-p *étale* quotient (as any such quotient must kill the μ_p we have inside $\mathcal{E}[p]$.)

Thus,

$$\begin{aligned}(\pi_2)_*\pi_1^*((\mathcal{E},\mathcal{P})_{/\overline{K}}) &= \sum_{\mathcal{C}} (\mathcal{E}/\mathcal{C}, \mathcal{P} \bmod \mathcal{C}) - \sum_{\mathcal{C}'} (\mathcal{E}'/\mathcal{C}', \mathcal{P}' \bmod \mathcal{C}') \\ &\in \mathrm{Pic}^0_{X_1(N)_{/\mathbf{Z}[\frac{1}{N}]}}(R)\end{aligned}$$

coincides with $(T_p)_*((\mathcal{E},\mathcal{P})-(\mathcal{E}',\mathcal{P}'))$ as both induce the same $\overline{K}$-point. Passing to closed fibers,

$$\begin{aligned}(T_p)_*((x)-(x')) = {} & (E/\mu_p, P \bmod \mu_p) + p(E/\underline{\mathbf{Z}/p\mathbf{Z}}', P \bmod \mathbf{Z}/p\mathbf{Z}) \\ & - (E'/\mu_p, P' \bmod \mu_p) + p(E'/\underline{\mathbf{Z}/p\mathbf{Z}}, P' \bmod \mathbf{Z}/p\mathbf{Z})\end{aligned}$$

where $E[p] \cong \mu_p \times \underline{\mathbf{Z}/p\mathbf{Z}}$ and $E'[p] \cong \mu_p \times \underline{\mathbf{Z}/p\mathbf{Z}}$ are the *canonical* splittings of the connected-étale sequence over the perfect field $\overline{\mathbf{F}}_p$.

Now consider the relative Frobenius morphism

$$F_{E/\overline{\mathbf{F}}_p} : E \to E^{(p)},$$

which sends O to O (and P to $P^{(p)}$) and so is a map of *elliptic curves* over $\overline{\mathbf{F}}_p$. Recall that in characteristic p, for any map of schemes $X \to S$ we define the relative Frobenius map $F_{X/S} : X \to X^{(p)}$ to be the unique S-map fitting into the diagram

$$\begin{array}{ccccc} X & \xrightarrow{F_{X/S}} & X^{(p)} & \longrightarrow & X \\ & \searrow & \downarrow & & \downarrow \\ & & S & \xrightarrow[F_S]{} & S \end{array} \qquad (F_X : X \to X)$$

where F_S, F_X are the absolute Frobenius maps. Since $E \to \operatorname{Spec} \overline{\mathbf{F}}_p$ is smooth of pure relative dimension 1, $F_{E/\overline{\mathbf{F}}_p}$ is finite flat of degree $p^1 = p$. It is bijective on points, so $\ker(F_{E/\overline{\mathbf{F}}_p})$ must be connected of order p.

The *only* such subgroup-scheme of E is $\mu_p \hookrightarrow E[p]$ by the *ordinariness*. Thus

$$E/\mu_p \cong E^{(p)}$$

is easily seen to take $P \bmod \mu_p$ to $P^{(p)}$.

Similarly, we have

$$E \xrightarrow[F_{E/\overline{\mathbf{F}}_p}]{} E^{(p)} \xrightarrow[F^{\vee}_{E/\mathbf{F}_p}]{} E \qquad (p : E \to E)$$

so $F^{\vee}_{E/\mathbf{F}_p}$ is étale of degree p and base extension by $\mathrm{Frob}^{-1} : \overline{\mathbf{F}}_p \to \overline{\mathbf{F}}_p$ gives

$$E^{(p^{-1})} \longrightarrow E \longrightarrow E^{(p^{-1})} \qquad (p)$$

$$P^{(p^{-1})} \longmapsto P \longmapsto p \cdot P^{(p^{-1})}.$$

As the second map in this composite is étale of degree p, we conclude

$$(E_{/\underline{\mathbf{Z}/p\mathbf{Z}}}, P \bmod \mathbf{Z}/p\mathbf{Z}) \cong (E^{(p^{-1})}, pP^{(p^{-1})}).$$

Thus, in $\mathrm{Pic}^0_{X_1(N)}(\overline{\mathbf{F}}_p)$,

$$\begin{aligned}(T_p)_*((x) - (x')) = {} & (E^{(p)}, P^{(p)}) + p \cdot (E^{(p^{-1})}, p \cdot P^{(p^{-1})}) \\ & - ((E')^{(p)}, (P')^{(p)}) - p \cdot ((E')^{(p^{-1})}, p \cdot (P')^{(p^{-1})})\end{aligned}$$

which we have seen is equal to $(F + \langle p \rangle_* F^{\vee})((x) - (x'))$.

□

CHAPTER 6

Appendix by Kevin Buzzard: A mod ℓ multiplicity one result

In this appendix, we explain how the ideas of [**46**] can be used to prove the following mild strengthening of the multiplicity one results in §9 of [**32**].

The setup is as follows. Let f be a normalised cuspidal eigenform of level N, and weight k, defined over $\overline{\mathbf{F}}_\ell$, with $\ell \nmid N$ and $2 \leq k \leq \ell + 1$. Let N^* denote N if $k = 2$, and $N\ell$ if $k > 2$. Let $J_{\mathbf{Q}}$ be the Jacobian of the curve $X_1(N^*)_{\mathbf{Q}}$, and let H denote the Hecke algebra in $\mathrm{End}(J_{\mathbf{Q}})$ generated over $\mathbf{Z}$ by T_p for all primes p, and all the Diamond operators of level N^*. It is well-known (for example by Proposition 9.3 of [**46**]) that there is a characteristic 0 normalised eigenform F in $S_2(\Gamma_1(N^*))$ lifting f. Let $\mathfrak{m}$ denote the maximal ideal of H associated to F (note that $\mathfrak{m}$ depends only on f and not on the choice of F), and let $\mathbf{F} = H/\mathfrak{m}$, which embeds naturally into $\overline{\mathbf{F}}_\ell$. Suppose the representation $\rho_f : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_\ell)$ associated to f is absolutely irreducible, and furthermore assume that if $k = \ell + 1$ then ρ_f is not isomorphic to a representation coming from a form of weight 2 and level N.

Theorem 6.1. *If ρ_f is ramified at ℓ, or if ρ_f is unramified at ℓ but $\rho_f(\mathrm{Frob}_\ell)$ is not a scalar matrix, then $J_{\mathbf{Q}}(\overline{\mathbf{Q}})[\mathfrak{m}]$ has $H/\mathfrak{m}$-dimension two, and hence is a model for (precisely one copy of) ρ_f.*

The motivation for putting ourselves in the setup above is that every absolutely irreducible modular mod ℓ representation has a twist coming from a modular form of level prime to ℓ and weight at most $\ell + 1$. In particular, every modular mod ℓ representation has a twist coming from a form satisfying the conditions of our setup. Furthermore, if f is as in our setup, then Theorems 2.5 and 2.6 of [**32**] tell us the precise structure of the restriction of ρ_f to D_ℓ, a decomposition group at ℓ. These results are explained in Section 2.2. Using them, it is easy to deduce

Corollary 6.2. *Let ρ be an absolutely irreducible modular mod ℓ representation, such that $\rho_f(D_\ell)$ is not contained within the scalars. Then some twist of ρ comes from a modular form satisfying the conditions of the theorem, and hence ρ is a multiplicity one representation in the sense of Remark 3.4.2.*

The theorem, commonly referred to as a "multiplicity one theorem", is a mild extension of results of Mazur, Ribet, Gross and Edixhoven. It was announced for $\ell = 2$ as Proposition 2.4 of [**9**] but the proof given there is not quite complete—in fact, the last line of the proof there is a little optimistic. I would hence like to thank Ribet and Stein for the opportunity to correct this oversight in [**9**].

Proof of Theorem. Firstly we observe that the only case not dealt with by Theorem 9.2 of [**32**] is the case when $k = \ell$ and ρ_f is unramified at ℓ, with $\rho_f(\mathrm{Frob}_\ell)$ a non-scalar matrix whose eigenvalues are equal. Moreover, using Theorems 2.5 and 2.6 of [**32**] we see that in this case f must be ordinary at ℓ. We are hence in a position to use the detailed construction of ρ_f given in §§11–12 of [**46**]. We will follow the conventions set up in the present paper for normalisations of Hecke operators, and so in particular the formulae below differ from the ones in [**46**] by a twist.

The maximal ideal $\mathfrak{m}$ of H associated to f gives rise as in (12.5) of [**46**] to an idempotent $e \in H_\ell := H \otimes_{\mathbf{Z}} \mathbf{Z}_l$, such that the completion $H_\mathfrak{m}$ of H at $\mathfrak{m}$ is just eH_ℓ. Let G denote $e(J_{\mathbf{Q}_\ell}[\ell^\infty])$, the part of the ℓ-divisible group of J which is associated to $\mathfrak{m}$. Then $H_\mathfrak{m}$ acts on G, and it is proved in Propositions 12.8 and 12.9 of [**46**] that there is an exact sequence of ℓ-divisible groups

$$0 \to G^0 \to G \to G^e \to 0$$

over $\mathbf{Q}_\ell$, which is $H_\mathfrak{m}$-stable. Let

$$0 \to T^0 \to T \to T^e \to 0$$

be the exact sequence of Tate modules of these groups. We now explain explicitly, following [**46**], how the group D_ℓ acts on these Tate modules.

If $k > 2$ then there is a Hecke operator U_ℓ in $H_\mathfrak{m}$, and we define $u = U_\ell$. If $k = 2$ then there is a Hecke operator T_ℓ in $H_\mathfrak{m}$ and because we are in the ordinary case we know that T_ℓ is a unit in $H_\mathfrak{m}$. We define u to be the unique root of the polynomial $X^2 - T_\ell X + \ell\langle\ell\rangle$ in $H_\mathfrak{m}$ which is a unit (u exists by an appropriate analogue of Hensel's lemma).

The calculations of Propositions 12.8 and 12.9 of [**46**] show that, under our conventions, the absolute Galois group D_ℓ of $\mathbf{Q}_\ell$ acts on T^e as $\lambda(u)$, where $\lambda(x)$ denotes the unramified character taking Frob_ℓ to x. Moreover, these propositions also tell us that D_ℓ acts on T^0 via the character $\chi_\ell\lambda(u^{-1}\langle\ell\rangle_N)\chi^{\ell-2}$, where χ_ℓ is the cyclotomic character and χ is the Teichmüller character. The key point is that this character takes values in $H^\times$.

The next key observation is that a standard argument on differentials, again contained in the proof of Propositions 12.8 and 12.9 of [**46**], shows that $G^e[\mathfrak{m}] = \mathfrak{m}^{-1}\ell T^e/\ell T^e$ has $H_\mathfrak{m}/\mathfrak{m}$-dimension 1 and that $G^0[\mathfrak{m}]$ has dimension $d^0 \geq 1$. (Note that the fact that $G^e[\mathfrak{m}]$ has dimension 1 implies, via some simple linear algebra, that the sequence $0 \to G^0[\mathfrak{m}] \to G[\mathfrak{m}] \to G^e[\mathfrak{m}] \to 0$ is exact, as asserted by Gross.) Furthermore, because we can identify $G^0[\mathfrak{m}]$ with $\mathfrak{m}^{-1}\ell T^0/\ell T^0$, we see that the action of D_ℓ on $G^0[\mathfrak{m}]$ is via a character which takes values in $(H/\mathfrak{m})^\times$. In particular, D_ℓ acts as scalars on $G^0[\mathfrak{m}]$.

Let us now assume that ρ_f is unramified at ℓ, and that $\rho_f(\mathrm{Frob}_\ell)$ is a non-diagonalisable matrix with eigenvalue $\alpha \in H/\mathfrak{m}$. Choose a model ρ for ρ_f defined over $\mathrm{GL}_2(H/\mathfrak{m})$. By the theorem of Boston, Lenstra and Ribet, we know that $G[\mathfrak{m}]$ is isomorphic to a direct sum of d copies of ρ, or more precisely, d copies of the restriction of ρ to D_ℓ. Here d is an integer satisfying $2d = d^0 + d^e$. Hence, if $G[\mathfrak{m}]^\alpha$ denotes the subspace of $G[\mathfrak{m}]$ where Frob_ℓ acts as α, then the $H/\mathfrak{m}$-dimension of $G[\mathfrak{m}]^\alpha$ is at most d. On the other hand, Frob_ℓ acts on $G[\mathfrak{m}]^0$ as a scalar, and hence this scalar must be α, and so we see $G[\mathfrak{m}]^0 \subseteq G[\mathfrak{m}]^\alpha$. Hence $d^0 \leq d = (d^0 + 1)/2$. We deduce that $d^0 \leq 1$ and hence $d^0 = d = 1$ and the theorem is proved. □

We remark that L. Kilford has found examples of mod 2 forms f of weight 2, such that ρ_f is unramified at 2 and $\rho_f(\mathrm{Frob}_2)$ is the identity, and where $J_{\mathbf{Q}}(\overline{\mathbf{Q}})[\mathfrak{m}]$ has $H/\mathfrak{m}$-dimension 4, and so one cannot hope to extend the theorem to this case. See Remark 3.6 for more details, or [**64**]. A detailed analysis of what is happening in this case, at least in the analogous setting of forms of weight 2 on $J_0(p)$, with p prime, has been undertaken by Emerton in [**39**]. In particular, Emerton proves that multiplicity one fails if and only if the analogue of the exact sequence $0 \to T^0 \to T \to T^e \to 0$ fails to split as a sequence of $H_{\mathfrak{m}}$-modules.

BIBLIOGRAPHY

1. A. Ash and G. Stevens, *Cohomology of arithmetic groups and congruences between systems of Hecke eigenvalues*, J. Reine Angew. Math. **365** (1986), 192–220.
2. A. O. L. Atkin and J. Lehner, *Hecke operators on* $\Gamma_0(m)$, Math. Ann. **185** (1970), 134–160.
3. B. J. Birch, *Cyclotomic fields and Kummer extensions*, Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), Thompson, Washington, D.C., 1967, pp. 85–93.
4. B. J. Birch and W. Kuyk (eds.), *Modular functions of one variable. IV*, Springer-Verlag, Berlin, 1975, Lecture Notes in Mathematics, Vol. 476.
5. S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Springer-Verlag, Berlin, 1990.
6. N. Boston, H. W. Lenstra, Jr., and K. A. Ribet, *Quotients of group rings arising from two-dimensional representations*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), no. 4, 323–328.
7. C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*, or Wild* 3*-adic exercises*, http://www.math.harvard.edu/HTML/Individuals/Richard_Taylor.html
8. S. Brueggeman, *The non-existence of certain Galois extensions unramified outside 5*, Journal of Number Theory **75** (1999), 47–52.
9. K. Buzzard, *On level-lowering for mod 2 representations*, to appear in Mathematics Research Letters.
10. K. Buzzard, M. Dickinson, N. Shepherd-Barron, and R. Taylor, *On icosahedral Artin representations*, in preparation.
11. H. Carayol, *Sur les représentations* ℓ*-adiques associées aux formes modulaires de Hilbert*, Ann. scient. Éc. Norm. Sup., 4^{eb} série **19** (1986), 409–468.
12. ———, *Sur les représentations galoisiennes modulo* ℓ *attachées aux formes modulaires*, Duke Math. J. **59** (1989), 785–801.
13. W. Casselman, *On representations of* GL_2 *and the arithmetic of modular curves*, Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972) (Berlin), Springer, 1973, pp. 107–141. Lecture Notes in Math., Vol. 349.
14. I. V. Čerednik, *Uniformization of algebraic curves by discrete arithmetic subgroups of* $\mathrm{PGL}_2(k_w)$ *with compact quotient spaces*, Mat. Sb. (N.S.) **100(142)**

(1976), no. 1, 59–88, 165.
15. R. F. Coleman, *Serre's conjecture: The Jugentraum of the 20th century*, Mat. Contemp. **6** (1994), 13–18, XII School of Algebra, Part I (Portuguese) (Diamantina, 1992).
16. R. F. Coleman and B. Edixhoven, *On the semi-simplicity of the U_p-operator on modular forms*, Math. Ann. **310** (1998), no. 1, 119–127.
17. R. F. Coleman and J. F. Voloch, *Companion forms and Kodaira-Spencer theory*, Invent. Math. **110** (1992), no. 2, 263–281.
18. B. Conrad, *Modular forms, cohomology, and the Ramanujan conjecture*, in preparation.
19. B. Conrad, F. Diamond, and R. Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations*, J. Amer. Math. Soc. **12** (1999), no. 2, 521–567.
20. J. E. Cremona, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, Cambridge, 1997.
21. C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, Interscience Publishers, a division of John Wiley & Sons, New York-London, 1962, Pure and Applied Mathematics, Vol. XI.
22. H. Darmon, *Serre's conjectures*, Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994), Amer. Math. Soc., Providence, RI, 1995, pp. 135–153.
23. H. Darmon, F. Diamond, and R. Taylor, *Fermat's last theorem*, Current developments in mathematics, 1995 (Cambridge, MA), Internat. Press, Cambridge, MA, 1994, pp. 1–154.
24. P. Deligne, *Formes modulaires et représentations ℓ-adiques.*, Sém. Bourbaki no. 355, 1968/69 (Berlin and New York), Springer-Verlag, 1971, Lecture Notes in Mathematics, Vol. 179, pp. 139–172.
25. P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972) (Berlin), Springer, 1973, pp. 143–316. Lecture Notes in Math., Vol. 349.
26. F. Diamond, *The refined conjecture of Serre*, Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993) (Cambridge, MA), Internat. Press, 1995, pp. 22–37.
27. F. Diamond and J. Im, *Modular forms and modular curves*, Seminar on Fermat's Last Theorem (Providence, RI), Amer. Math. Soc., 1995, pp. 39–133.
28. M. Dickinson, *On the modularity of certain 2-adic galois representations*, Harvard Ph.D. thesis (2000).
29. D. Doud, *S_4 and $\tilde{S}_4$ extensions of $\mathbf{Q}$ ramified at only one prime*, J. Number Theory **75** (1999), no. 2, 185–197.
30. V. G. Drinfeld, *Coverings of p-adic symmetric domains*, Funkcional. Anal. i Prilov zen. **10** (1976), no. 2, 29–40.
31. B. Edixhoven, *L'action de l'algèbre de Hecke sur les groupes de composantes des jacobiennes des courbes modulaires est "Eisenstein"*, Astérisque (1991), no. 196–197, 7–8, 159–170 (1992), Courbes modulaires et courbes de Shimura (Orsay, 1987/1988).
32. ———, *The weight in Serre's conjectures on modular forms*, Invent. Math. **109** (1992), no. 3, 563–594.

33. B. Edixhoven, *Le rôle de la conjecture de Serre dans la démonstration du théorème de Fermat*, Gaz. Math. (1995), no. 66, 25–41.
34. ———, *Erratum and addendum: "The role of Serre's conjecture in the proof of Fermat's theorem"*, Gaz. Math. (1996), no. 67, 19.
35. ———, *Serre's conjecture*, Modular forms and Fermat's last theorem (Boston, MA, 1995) (New York), Springer, 1997, pp. 209–242.
36. M. Eichler, *Quadratische Formen und Modulfunktionen*, Acta Arith. **4** (1958), 217–239.
37. D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Springer-Verlag, New York, 1995.
38. D. Eisenbud and J. Harris, *Schemes, The language of modern algebraic geometry*, Springer-Verlag, Berlin, Graduate Texts in Mathematics, Vol. 197.
39. M. Emerton, *Supersingular elliptic curves, theta series and weight two modular forms*, preprint.
40. G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349–366.
41. G. Faltings and B. W. Jordan, *Crystalline cohomology and* $\mathrm{GL}(2, \mathbf{Q})$, Israel J. Math. **90** (1995), no. 1-3, 1–66.
42. G. Frey, *Links between stable elliptic curves and certain Diophantine equations*, Ann. Univ. Sarav. Ser. Math. **1** (1986), no. 1, iv+40.
43. ———, *Links between solutions of* $A - B = C$ *and elliptic curves*, Number theory (Ulm, 1987), Springer, New York, 1989, pp. 31–62.
44. A. Fröhlich, *Local fields*, Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), Thompson, Washington, D.C., 1967, pp. 1–41.
45. K. Fujiwara, *Level optimization in the totally real case*, in preparation (1999).
46. B. H. Gross, *A tameness criterion for Galois representations associated to modular forms (mod p)*, Duke Math. J. **61** (1990), no. 2, 445–517.
47. R. Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52.
48. Y. Hellegouarch, *Invitation aux mathématiques de Fermat-Wiles*, Masson, Paris, 1997.
49. H. Hida, *Galois representations into* $\mathrm{GL}_2(\mathbf{Z}_p[[X]])$ *attached to ordinary cusp forms*, Invent. Math. **85** (1986), no. 3, 545–613.
50. ———, *Iwasawa modules attached to congruences of cusp forms*, Ann. Sci. École Norm. Sup. (4) **19** (1986), no. 2, 231–273.
51. H. Jacquet and R. P. Langlands, *Automorphic forms on* $\mathrm{GL}(2)$, Springer-Verlag, Berlin, 1970, Lecture Notes in Mathematics, Vol. 114.
52. F. Jarvis, *On Galois representations associated to Hilbert modular forms*, J. Reine Angew. Math. **491** (1997), 199–216.
53. ———, *Level lowering for modular mod* ℓ *representations over totally real fields*, Math. Ann. **313** (1999), no. 1, 141–160.
54. ———, *Mazur's principle for totally real fields of odd degree*, Compositio Math. **116** (1999), no. 1, 39–79.
55. N. Jochnowitz, *A study of the local components of the Hecke algebra mod* ℓ, Trans. Amer. Math. Soc. **270** (1982), no. 1, 253–267.
56. ———, *The index of the Hecke ring,* T_k*, in the ring of integers of* $T_k \otimes \mathbf{Q}$, Duke Math. J. **46** (1979), no. 4, 861–869.

57. B. W. Jordan and R. Livné, *Conjecture "epsilon" for weight $k > 2$*, Bull. Amer. Math. Soc. (N.S.) **21** (1989), no. 1, 51–56.
58. K. Joshi, *Remarks on methods of Fontaine and Faltings*, Internat. Math. Res. Notices **1999**, no. 22, 1199–1209.
59. N. M. Katz, *p-adic properties of modular schemes and modular forms*, Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972) (Berlin), Springer, 1973, pp. 69–190. Lecture Notes in Mathematics, Vol. 350.
60. ———, *Higher congruences between modular forms*, Ann. of Math. (2) **101** (1975), 332–367.
61. ———, *A result on modular forms in characteristic p*, Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976) (Berlin), Springer, 1977, pp. 53–61. Lecture Notes in Math., Vol. 601.
62. N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Princeton University Press, Princeton, N.J., 1985.
63. C. Khare, *Multiplicities of mod p Galois representations*, Manuscripta Math. **95** (1998), no. 2, 181–188.
64. L. J. P. Kilford, *Some examples of non-Gorenstein Hecke algebras associated to modular forms*, in preparation.
65. A. W. Knapp, *Elliptic curves*, Princeton University Press, Princeton, NJ, 1992.
66. S. Lang, *Introduction to modular forms*, Springer-Verlag, Berlin, 1995, With appendixes by D. Zagier and Walter Feit, Corrected reprint of the 1976 original.
67. R. P. Langlands, *Modular forms and ℓ-adic representations*, Proceedings of the International Summer School, University of Antwerp, RUCA, July 17–August 3, 1972 (Berlin) (P. Deligne and W. Kuyk, eds.), Springer, 1973, pp. 361–500. Lecture Notes in Math., Vol. 349.
68. ———, *Base change for* GL(2), Princeton University Press, Princeton, N.J., 1980.
69. W-C. Li, *Newforms and functional equations*, Math. Ann. **212** (1975), 285–315.
70. R. Livné, *On the conductors of mod ℓ Galois representations coming from modular forms*, J. Number Theory **31** (1989), no. 2, 133–141.
71. B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes Études Sci. Publ. Math. (1977), no. 47, 33–186 (1978).
72. B. Mazur and K. A. Ribet, *Two-dimensional representations in the arithmetic of modular curves*, Astérisque (1991), no. 196-197, 6, 215–255 (1992), Courbes modulaires et courbes de Shimura (Orsay, 1987/1988).
73. L. Merel, *Universal Fourier expansions of modular forms*, On Artin's conjecture for odd 2-dimensional representations (Berlin), Springer, 1994, pp. 59–94.
74. J. S. Milne, *Abelian varieties*, Arithmetic geometry (Storrs, Conn., 1984), Springer, New York, 1986, pp. 103–150.
75. T. Miyake, *Modular forms*, Springer-Verlag, Berlin, 1989, Translated from the Japanese by Yoshitaka Maeda.
76. H. Moon, *Finiteness results on certain mod p Galois representations*, to appear in J. Number Theory.

77. D. Mumford, *Abelian varieties*, Published for the Tata Institute of Fundamental Research, Bombay, 1970, Tata Institute of Fundamental Research Studies in Mathematics, No. 5.
78. C. Queen, *The existence of p-adic Abelian L-functions*, Number theory and algebra (New York), Academic Press, 1977, pp. 263–288.
79. A. Raji, *On the levels of modular mod ℓ Galois representations of totally real fields*, Princeton University Ph.D. thesis, 1998.
80. R. Ramakrishna, *Lifting Galois representations*, Invent. Math. **138** (1999), no. 3, 537–562.
81. ———, *Deforming Galois representations and the conjectures of Serre and Fontaine-Mazur*, preprint, ftp://math.cornell.edu/pub/ravi (2000).
82. M. Raynaud, *Spécialisation du foncteur de Picard*, Inst. Hautes Études Sci. Publ. Math. No. **38** (1970), 27–76.
83. K. A. Ribet, *From the Taniyama-Shimura conjecture to Fermat's last theorem*, Ann. Fac. Sci. Toulouse Math. (5) **11** (1990), no. 1, 116–139.
84. ———, *On modular representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), no. 2, 431–476.
85. ———, *Raising the levels of modular representations*, Séminaire de Théorie des Nombres, Paris 1987–88, Birkhäuser Boston, Boston, MA, 1990, pp. 259–271.
86. ———, *Lowering the levels of modular representations without multiplicity one*, International Mathematics Research Notices (1991), 15–19.
87. ———, *Report on mod ℓ representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Motives (Seattle, WA, 1991), Amer. Math. Soc., Providence, RI, 1994, pp. 639–676.
88. A. Robert, *Elliptic curves*, Springer-Verlag, Berlin, 1973, Notes from postgraduate lectures given in Lausanne 1971/72, Lecture Notes in Mathematics, Vol. 326.
89. T. Saito, *Modular forms and p-adic Hodge theory*, Invent. Math. **129** (1997), 607–620.
90. I. Schur, *Arithmetische Untersuchungen über endliche Gruppen linearer Substitutionen*, Sitz. Pr. Akad. Wiss. (1906), 164–184, Gesam. Abhl., **I**, 177–197, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1973.
91. J-P. Serre, *Groupes de Lie l-adiques attachés aux courbes elliptiques*, Les Tendances Géom. en Algébre et Théorie des Nombres, Éditions du Centre National de la Recherche Scientifique, Paris, 1966, pp. 239–256 (= Collected Papers **70**).
92. ———, *Une interprétation des congruences relatives à la fonction τ de Ramanujan*, Séminaire Delange-Pisot-Poitou n^o **14** (1967–68) (= C.P. **80**).
93. ———, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), no. 4, 259–331 (= C.P. **94**).
94. ———, *Congruences et formes modulaires [d'après H. P. F. Swinnerton-Dyer]*, Séminaire Bourbaki, 24e année (1971/1972), Exp. No. 416 (Berlin), Springer, 1973, pp. 319–338. Lecture Notes in Math., Vol. 317 (= C.P. **95**).
95. ———, *Formes modulaires et fonctions zêta p-adiques*, Proceedings of the International Summer School, University of Antwerp, RUCA, July 17–August 3, 1972 (Berlin), Springer, 1973, pp. 191–268. Lecture Notes in Math., Vol. 350 (= C.P. **97**).

96. ———, *A Course in Arithmetic*, Springer-Verlag, New York, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7.
97. ———, *Valeurs propres des opérateurs de Hecke modulo ℓ*, Astérisque **24–25** (1975), 109–117 (= C.P. **104**).
98. ———, *Divisibilité de certaines fonctions arithmétiques*, Enseign. Math. (2) **22** (1976), no. 3-4, 227–260 (= C.P. **108**).
99. ———, *Linear representations of finite groups*, Springer-Verlag, New York, 1977, Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.
100. ———, *Local fields*, Springer-Verlag, New York, 1979, Translated from the French by Marvin Jay Greenberg.
101. ———, *Lettre à J.-F. Mestre*, Current trends in arithmetical algebraic geometry (Arcata, Calif., 1985), Amer. Math. Soc., Providence, RI, 1987, pp. 263–268 (= C.P. **142**).
102. ———, *Sur les représentations modulaires de degré* 2 *de* Gal($\overline{\mathbf{Q}}/\mathbf{Q}$), Duke Math. J. **54** (1987), no. 1, 179–230 (= C.P. **143**).
103. ———, *Travaux de Wiles (et Taylor, ...)*, Partie I, Séminaire Bourbaki, **803** (1995) (= C.P. **168**).
104. J-P. Serre and J. T. Tate, *Good reduction of abelian varieties*, Ann. of Math. (2) **88** (1968), 492–517 (= C.P. **79**).
105. N. I. Shepherd-Barron and R. Taylor, *Mod 2 and mod 5 icosahedral representations*, J. Amer. Math. Soc. **10** (1997), no. 2, 283–298.
106. H. Shimizu, *On zeta functions of quaternion algebras*, Ann. of Math. (2) **81** (1965), 166–193.
107. G. Shimura, *A reciprocity law in non-solvable extensions*, J. Reine Angew. Math. **221** (1966), 209–220.
108. ———, *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, Princeton, NJ, 1994, Reprint of the 1971 original, Kan Memorial Lectures, 1.
109. J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.
110. ———, *Advanced topics in the arithmetic of elliptic curves*, Springer-Verlag, New York, 1994.
111. C. M. Skinner and A. J. Wiles, *Ordinary representations and modular forms*, Proc. Nat. Acad. Sci. U.S.A. **94** (1997), no. 20, 10520–10527.
112. H. P. F. Swinnerton-Dyer, *On ℓ-adic representations and congruences for coefficients of modular forms*, Proceedings of the International Summer School, University of Antwerp, RUCA, July 17–August 3, 1972 (Berlin), Springer, 1973, pp. 1–55. Lecture Notes in Math., Vol. 350.
113. J. T. Tate, *The non-existence of certain Galois extensions of* $\mathbf{Q}$ *unramified outside* 2, Contemporary Math. **174** (1994), 153–156.
114. R. Taylor and A. J. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2) **141** (1995), no. 3, 553–572.
115. J. Tunnell, *Artin's conjecture for representations of octahedral type*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), no. 2, 173–175.
116. J.-L. Waldspurger, *Quelques propriétés arithmétiques de certaines formes automorphes sur* GL(2), Compositio Math. **54** (1985), no. 2, 121–171.

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

INDEX

Italic page numbers are used to indicate pages with important information about the entry, while page numbers in normal type indicate a textual reference.