How can we construct abelian Galois extensions of basic number fields?

Barry Mazur

July 5, 2010

1 Introduction

Recall the almost tautological—but very useful—way we have of dealing with an irreducible polynomial over a field. For example, in the case of $X^5 - X + 1$, irreducible over the rational field $\mathbf{Q}$, we happily adjoin a root to our base field by merely forming $L :=$ the quotient of the polynomial ring $\mathbf{Q}[X]$ modulo the ideal generated by $X^5 - X + 1$. Then L is a field (of algebraic numbers) and the image of X is indeed a root of $X^5 - X + 1$ in L.

This method is serviceable, as far as it goes[1], but sometimes the field extensions that we are interested in, the field extensions that we expect—thanks to some heuristic or other—*should exist*, would not be readily constructible this way, nor—once constructed—would they be understandable, nor treatable, this way: i.e., in terms of polynomials whose roots generate them, even if those polynomials were readily available. Such is the case, for the most part, for the abelian extensions alluded to in the title above.

This article is based on a talk I gave entitled *Construction of abelian extensions following Ken Ribet* at the 60th Birthday Conference for Ken Ribet (held at the University of California at Berkeley and MSRI June 28 - July 2, 2008) [2]. My mission was to focus on Ribet's method of construction of abelian Galois extensions of cyclotomic fields—one of Ken's great early achievements—and to hint at the vast influence this work has had in the later development of our subject[3].

The central topic of this article is the theorem referred to nowadays simply as *The Herbrand-Ribet Theorem.*

The first thing to know about this theorem is that Ken's 11-page paper [**?**] from which much of this mathematics stems is as worth reading today as it was over three decades ago, and is eminently

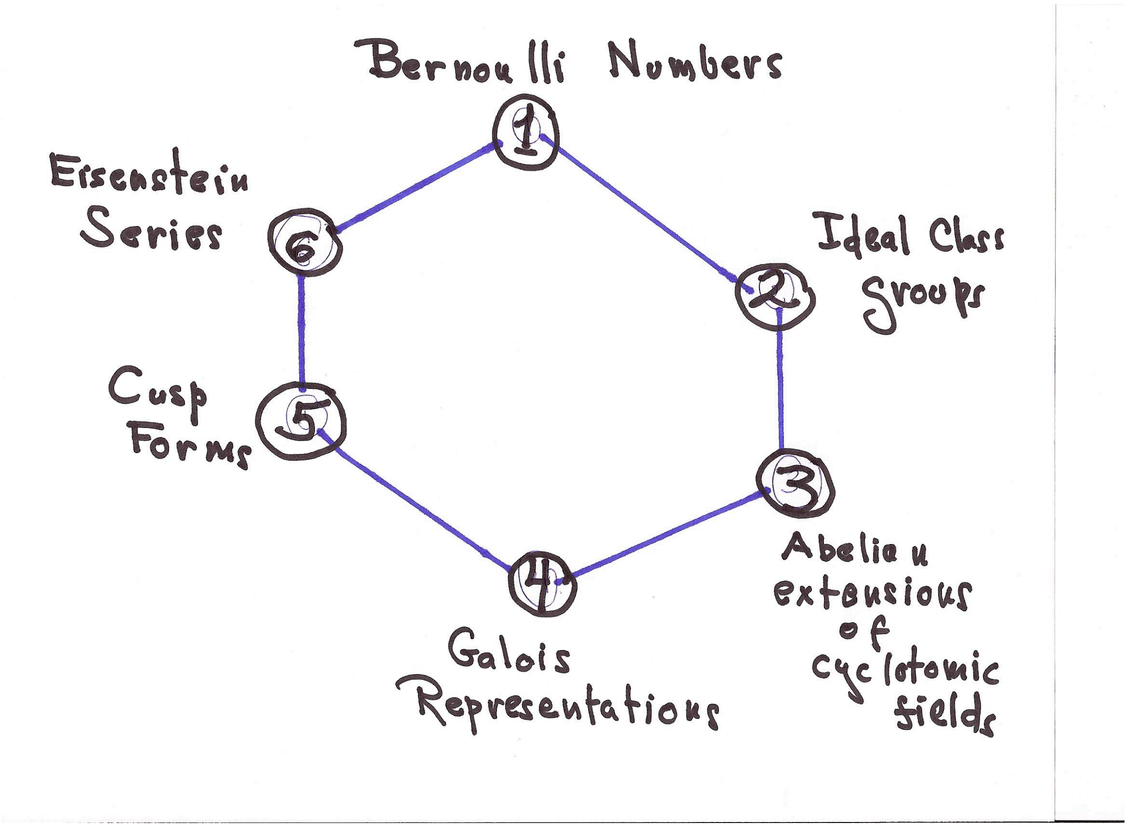
[1] and on occasion, works quite nicely, as in the construction of cyclotomic fields in section **??** below

[2] I want to thank Joel Belläiche, Gaetan Chenevier, Ralph Greenberg, Michael Harris, and Eric Urban who have helped me with comments and corrections regarding early drafts of this article.

[3] It was fun to do that, and it was a particular pleasure to me; one of the joys of doing mathematics is that you get people like Ken Ribet as friends and colleagues.

readable. For this reason, one aim of this article is not to give a proof of the theorem (I don't) but rather to try to explain why the ideas behind the theorem have played such an inspirational role in the subject, and why they will continue to do so.

The second thing to know is that the Herbrand-Ribet Theorem for a prime number p concerns *six* slightly different facets of number theory, specifically as they are related to p, and weaves them together in a striking way.



This article is divided into five parts and an appendix.

Part I is a general introduction to the hexagon above and a discussion of how its vertices are linked together. We will do this by circumnavigating it three times:

- In the first round we merely give some preliminary hints about some of these facets of number theory and their connection to each other.
- The second circuit will be a more precise run-through using the prime $p = 691$ as an example. Here we will be highlighting six specific objects (or specific computations) occurring in the parts of number theory that correspond to each of our vertices. The fact that these six phenomena are related and that the linked chain that they form provides us with a powerful way of understanding each of them—and indeed constructing some of them—offers yet more evidence if we ever needed it that mathematics is an indivisible whole.

- After some discussion of background material we do the third lap around the hexagon, to formulate the Herbrand-Ribet Theorem for the general prime number p.

I hope that people who wish to get the general flavor of the number theory involved in this hexagon will be able to do so whether or not they have the background for some of the more detailed issues discussed in the later parts of this article.

Part II discusses results that give us interesting Galois representations are (a) managed efficiently by knowledge of Frobenius eigenvalues and (b) arise from Algebraic Geometry.

Part III takes up these ideas more explicitly and discusses some of the issues that relate to what I call Ribet's "wrench."

Part IV deals with the taxonomy of different types of Galois representations.

Part V hints at how the 'Ribet philosophy,' taken broadly, is continuing to inspire current work in the area.

Part I: About the Herbrand-Ribet Theorem

2 Cyclotomic Number Fields and their arithmetic

To launch into my topic, the "basic number fields" referred to in the title are the *cyclotomic number fields.* A cyclotomic number field is a field generated over the rational field $\mathbf{Q}$ by the adjunction of a primitive N-th root of unity, for some N. For example, we can view this field as the subfield of the field of complex numbers generated by $e^{2\pi i/N}$.

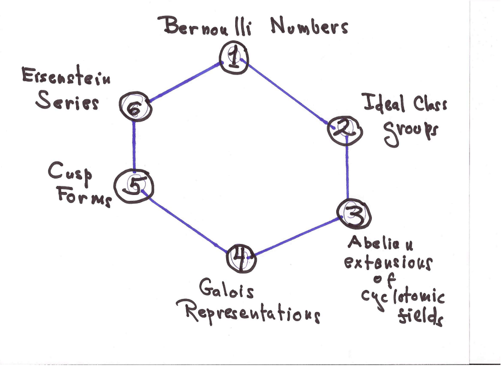
The "first two" of these cyclotomic number fields—i.e. $\mathbf{Q}(e^{2\pi i/N})$ for $N = 3, 4$—are thoroughly familiar to many mathematicians: they are the quadratic number fields $\mathbf{Q}(\sqrt{-3})$ and $\mathbf{Q}(\sqrt{-1})$ sitting nicely in the complex plane and have the property that their rings of integers represent elegantly symmetric lattices in the complex plane: for $N = 3$ one gets a hexagonal lattice; for $N = 4$, where the ring of integers in the cyclotomic field $\mathbf{Q}(\sqrt{-1})$ is the *ring of gaussian integers* $\{a + ib \mid a, b \in \mathbf{Z}\}$, the corresponding lattice is the square lattice. The "next" cyclotomic field in turn, i.e. for $N = 5$ also has had a great role to play in our subject in that it is a quadratic extension of the (quadratic) number field generated over $\mathbf{Q}$ by the "golden mean;" it is the field relevant for the classical construction of the regular pentagon.

But perhaps I shouldn't be going—one by one—through the list of these cyclotomic number fields, for the totality of them have played a crucial role in the development of mathematics in general, and arithmetic in particular. Their crucial importance was certainly recognized by Gauss, where the field extension $\mathbf{Q}(e^{2\pi i/N})/\mathbf{Q}$ was seen to be related to the construction of the regular N-gon, and to be a key to the fuller understanding of all quadratic number fields. By the latter part of the nineteenth century, thanks to the work of Kummer, it was known that the fine arithmetic features of these fields gave powerful methods to view (systematically) abelian extensions of number fields as generated by radicals, and to approach Fermat's Last Theorem for regular[4] prime exponents. The modern arithmetic of cyclotomic fields per se is enriched by the theory of Iwasawa who surmised a certain profound connection between the structure of the ideal class groups[5]—or equivalently—of abelian everywhere unramified extensions—of cyclotomic number fields and a p-adic interpolation of analytic number theory; and more specifically, the p-adic version of Dirichlet L-functions as constructed by Kubota and Leopoldt[6].

In view of all this, one can see why Serge Lang once referred to the arithmetic theory of cyclotomic fields as the "backbone of algebraic number theory."

3 Six facets of Number Theory

Here again are labels for the six topics:

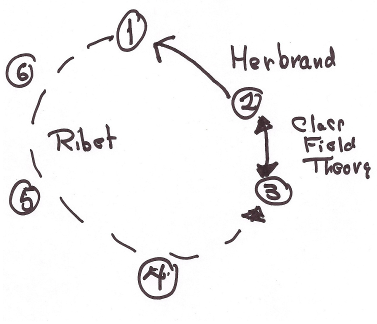


and here is a cartoon of how they get connected to each other:

[4] and some irregular

[5] —more specifically—for p a prime number, the p-primary components of the ideal class groups of cyclotomic fields obtained by the extraction of N roots of unity where N is a power of p.

[6] The seeds for such a program occurs already in the latter part of the 19th century in the work of E. Kummer, and the foundations for this program occurs in the work of Hensel.



In a sentence, Ken Ribet managed to pass from

- properties of Bernoulli numbers to
- note a consequence about Eisenstein series which
- allows him to construct certain cuspidal modular forms related to these Eisenstein series, and in turn to
- construct, using these cuspforms, a certain homomorphism of the automorphism group of $\bar{\mathbf{Q}}$, the field of algebraic numbers, to the group of upper triangular matrices in $\mathrm{GL}_2(\mathbf{F}_p)$ so that he may then
- form the number field $L \subset \bar{\mathbf{Q}}$ consisting of all the algebraic numbers fixed by the subgroup of automorphisms comprising the kernel of this homomorphism,
- this L being *the* abelian extension of the cyclotomic field that was *desired to be* constructed and whose properties were to be verified,

thereby completing the circuit of the second figure.

4 A few words about some of the stations and how they are connected to each other

4.1 Abelian extensions of number fields and ideal class groups

That there is a connection between these two arithmetic objects is seen already in Gauss's *Disquisitiones Arithmeticae* and has been a theme threading through work of Dirichlet, Dedekind, Kummer, Hilbert, leading to the more complete Class Field Theory of Takagi, Artin, and Chevalley.

For K a number field, Class Field Theory offers us a construction of the category of abelian Galois extensions of K in relatively concrete terms[7]. An important special case of this is the isomorphism that Class Field Theory provides, between $\mathrm{Gal}(H/K)$, the Galois group of the maximal abelian *everywhere unramified* Galois extension of K, and the classical ideal class group, $\mathcal{C}l(K)$, of K. A consequence of this isomorphism for us is that the problem of constructing quotient groups of the ideal class group of K with certain properties translates to constructing abelian unramified extensions of K with the corresponding properties. That is, it provides us with a link between stations $\boxed{2}$ and $\boxed{3}$.

4.2 "Constructing" Galois extensions and Galois representations

Let us review the standard way of studying the cyclotomic field L obtained by adjoining a primitive N-th root of unity to the rational field. This number field can be thought of as *the* splitting field of the polynomial $X^N - 1$ over $\mathbf{Q}$. If μ_N denotes the group (under multiplication) of all N-th roots of unity in this splitting field we have that μ_N is a cyclic group of order N, and it is the set of all roots of $X^N - 1$. More precisely,

$$X^N - 1 \;=\; \prod_{\zeta \in \mu_N} (X - \zeta).$$

If $G := \mathrm{Gal}(L/\mathbf{Q})$ denotes the Galois group of L over $\mathbf{Q}$, then G preserves the set of roots of $X^N - 1$, and thus can be thought of as a subgroup of the group of permutations of the set μ_N. Even better, since the action of G commutes with multiplication, G acts as a group of automorphisms of the *cyclic group* μ_N. So we get a natural imbedding

$$\mathrm{Gal}(L/\mathbf{Q}) \;\hookrightarrow\; \mathrm{Aut}_{\mathrm{gp}}(\mu_N) \;=\; (\mathbf{Z}/N\mathbf{Z})^* = \mathrm{GL}_1(\mathbf{Z}/N\mathbf{Z}).$$

Note that the middle equality is given by the canonical isomorphism[8] of $(\mathbf{Z}/N\mathbf{Z})^*$—the group of units in the ring $\mathbf{Z}/N\mathbf{Z}$—with the group of automorphisms of the cyclic group μ_N. Thanks to a classical theorem of Gauss the first inclusion is an isomorphism, giving us one of the most venerable surjective Galois representations in the history of the subject:

$$\omega_N : \mathrm{Gal}(L/\mathbf{Q}) \;\xrightarrow{\simeq}\; \mathrm{GL}_1(\mathbf{Z}/N\mathbf{Z}),$$

[7] Class Field Theory provides us with an isomorphism between the abelianization of G_K and the profinite group of connected components of the idele class group of K. It is no accident, though, that *abelian* Galois extensions are more amenable to detailed study than more general Galois extensions, and this is not only because abelian groups are easier to study than more general groups. Just as, in algebraic topology one must specify a base point to explicitly define the fundamental group of a connected space, but one needn't do this to define its homology group, so too—with a field K one must specify a separable algebraic closure, K^{sep}, of K to explicitly define the full Galois group $G_K := \mathrm{Gal}(K^{\mathrm{sep}}/K)$. Without such a specification, G_K is only defined "up to conjugation" (which is why it is particular fitting to study the structure of G_K via its linear representations—which are themselves only defined up to conjugation) but one has no need to do this to define the abelianization of G_K,

$$G_K^{\mathrm{ab}} := G_K/G'_K.$$

(here G'_K is the closed normal subgroup generated by commutators) whose quotients by closed subgroups of finite index provide us with the Galois groups of all abelian extensions of K.

[8] Associate to $\alpha \in (\mathbf{Z}/N\mathbf{Z})^*$ the automorphism $\zeta \mapsto \zeta^\alpha$ for $\zeta \in \mu_N$.

a faithful degree one representation of $\mathrm{Gal}(L/\mathbf{Q})$ over the ring $\mathbf{Z}/N\mathbf{Z}$.

It is standard, nowadays, to construct Galois extensions over a field K by finding natural continuous actions of $G_K := \mathrm{Gal}(\bar{K}/K)$ on vector spaces or on modules over rings—these coming to us usually as cohomology modules. Here $\bar{K}$ is an algebraic closure of K.

For example, if a module H, free of rank d over a finite ring R, is endowed with a with a continuous R-linear G_K-action we would then get a representation

$$\rho : G_K \to \mathrm{Aut}_R(H) \simeq \mathrm{GL}_d(R).$$

Call these things *Galois representations.* Since we have assumed that our ring R is finite, the image of ρ is finite, there is a unique finite-degree extension L/K with $L \subset \bar{K}$ such that ρ factors through the injective homomorphism

$$G_K \to \mathrm{Gal}(L/K) \hookrightarrow \mathrm{GL}_d(R).$$

It is in this sense that our Galois representation ρ has "constructed" L/K (together with a faithful degree d representation of $\mathrm{Gal}(L/K)$ over the ring R).

When we get to station $\boxed{4}$ of the above diagram, we'll be touching on such constructions again; in particular, in section **??** below.

4.3 Abelian extensions and two-dimensional Galois representations

Our desired abelian extensions L (as in station $\boxed{3}$ in the diagram above) will in fact be cyclic of order p over the cyclotomic number field $\mathbf{Q}(e^{2\pi i/p})$, but decidedly *non*-abelian over $\mathbf{Q}$, the field of rational numbers. It is natural, then (although it might seem odd—at first glance) that the systematic way of constructing these extensions is to first construct a non-abelian (two-dimensional, in fact) indecomposable representation of $G_{\mathbf{Q}}$, the Galois group of an algebraic closure of $\mathbf{Q}$, into $\mathrm{GL}_2(\mathbf{F}_p)$ and to get the required extension field as the field cut out by this degree two representation over $\mathbf{F}_p$, in the manner discussed in subsection **??** above. This type of argument will be giving us a link between stations $\boxed{3}$ and $\boxed{4}$.

4.4 Modular forms and the two-dimensional Galois representations that are associated to them

Cuspforms[9] come into play, for they are a very convenient source of *irreducible* continuous two-dimensional $G_{\mathbf{Q}}$ representations over the field of p-adic numbers, $\mathbf{Q}_p$, and over finite degree extensions of $\mathbf{Q}_p$, for any prime p. For some introductory discussion about this see sections **??** and **??** below. For a slightly more descriptive discussion of the passage from cuspforms to Galois representations via Deligne's Theorem see section **??**.

[9] More specifically, in this article we will only be interested eigenforms for the Hecke operators. For a fine introduction to this material and its connection to L-functions and modular curves, see Rohrlich's [**?**].

By suitably reducing these $G_{\mathbf{Q}}$ representations to characteristic p, one gets a supply of continuous two-dimensional $G_{\mathbf{Q}}$ representations over finite fields k of characteristic p; all the representations so obtained have the further property that the complex conjugation involution does not act as a scalar. The famous conjecture of Serre ([**?**]), recently proved by Khare and Wintenberger ([**?**]) asserts that *every irreducible* continuous two-dimensional $G_{\mathbf{Q}}$ representation over a finite field in which complex conjugation involution does not act as a scalar comes, in this manner, from a cuspform.

4.5 The Ribet wrench: first encounter

Ribet's goal is to find certain reducible-but-indecomposable two-dimensional Galois representations over finite fields, say $\mathbf{F}_p$. By choosing a suitable basis of the two-dimensional $\mathbf{F}_p$-vector space, such a representation can be given by a homomorphism

$$G_{\mathbf{Q}} \to \text{Upper triangular matrices } \subset \mathrm{GL}_2(\mathbf{F}_p),$$

$$g \mapsto \begin{pmatrix} \chi_1(g) & b(g) \\ 0 & \chi_2(g) \end{pmatrix}$$

where χ_1 and χ_2 are characters of the Galois group with values in $\mathbf{F}_p^*$. More succinctly we evoke such a representation by the following matrix picture:

$$\begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix}.$$

The intersection of the kernels of these two characters define–in the usual manner of Galois theory—an abelian field extension M of the rational field $\mathbf{Q}$; that is, M is the smallest subfield of the algebraic closure $\bar{\mathbf{Q}}$ such that the restriction of those characters to the subgroup $\mathrm{Gal}(\bar{\mathbf{Q}}/M) \subset \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ is trivial. Thus M is the smallest field extension for which above representation, when restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/M)$, has the shape:

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

.

In the theory we are presenting, the field M is usually the p-cyclotomic field $\mathbf{Q}(e^{2\pi i/p})$.

If the representation restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(e^{2\pi i/p}))$ *is nontrivial* then the image of $\mathrm{Gal}(\bar{\mathbf{Q}}/M)$ in $\mathrm{GL}_2(\mathbf{F}_p)$ is isomorphic to the group of unipotent upper triangular matrices displayed above, and hence is a cyclic group of order p cutting out a cyclic Galois extension $N/\mathbf{Q}(e^{2\pi i/p})$. So if we have an explicit understanding of the initial two-dimensional representation of $G_{\mathbf{Q}}$ we have, in similar explicit terms, "constructed" the cyclic degree p extension N of the p-cyclotomic field.

Ribet achieves this aim by

- initially finding an irreducible two-dimensional Galois representation over a field of characteristic 0,—say over the field $\mathbf{Q}_p$ p-adic numbers; that is, a continuous homomorphism

$$\rho : G_{\mathbf{Q}} \to \mathrm{Aut}_{\mathbf{Q}_p}(V) \cong \mathrm{GL}_2(\mathbf{Q}_p)$$

where V is a two-dimensional $\mathbf{Q}_p$-vector space, and the action of $G_{\mathbf{Q}}$ preserves no line.

- He then appropriately reduces this Galois representation modulo p. To do this, though, he must *choose* a $\mathbf{Z}_p$-lattice $\Omega \subset V$ that is stabilized by the action of $G_{\mathbf{Q}}$ (such a lattice always exists) and then pass to the action of $G_{\mathbf{Q}}$ on $\bar{V} := \Omega/p\Omega$.

  In this manner you get a representation mod p, which we will call the *residual representation* obtained from the lattice Ω:

$$\bar{\rho} : G_{\mathbf{Q}} \to \mathrm{Aut}_{\mathbf{Q}_p}(\bar{V}) \cong \mathrm{GL}_2(\mathbf{F}_p).$$

  A priori, the equivalence class of the representation $\bar{\rho}$ depends on the choice of lattice Ω. Now, there are many such lattices Ω that are stabilized by the $G_{\mathbf{Q}}$-action so we need to discuss how *many* inequivalent representations $\bar{\rho}$ might be obtained by varying the choice of Ω. For a start, multiplying any such lattice by a nonzero scalar provides another stable lattice, but it is evident that the corresponding residual representations of lattices that differ by scalar change are isomorphic.

  There are, however, situations where the residual representation may change (a bit) depending upon the lattice Ω that is chosen. What is *independent* of the lattice chosen is the *semi-simplification* of the residual representation $\bar{\rho}$. In particular, if the residual representation associated to one lattice stabilized by the action of $G_{\mathbf{Q}}$ is reducible, then it is so for all such lattices, and the two one-dimensional representations–i.e. characters— into which the two dimensional residual representation decompose is independent of the lattice chosen. But the two-dimensional reducible residual representation $\bar{\rho}$ itself may indeed depend upon the lattice.

- The key to Ribet's construction is to start with representations ρ that are irreducible, and yet have residual representations that are reducible; that is, in the terminology of the discussion above, the residual representations have matrix pictures looking like:

$$\begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix}$$

  and where, again as in the discussion above, one has constructed a cyclic p-extension if the corresponding residual representation, when restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/M)$, that has the shape:

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

  is *nontrivial.*

- In the above situation one can *always* change the lattice Ω suitably[10] to guarantee that the above representation is nontrivial.

[10] I envision it as a kind of “wrenching action”

- The *irreducibility* of his initial representation is crucial to his method. In its later manifestations this type of strategy can be framed as follows: you are looking for a “somewhat degenerate object” (e.g., a reducible representation) and you hope to get it as a “degenerate” member of, say, a one-parameter family of nondegenerate objects (e.g., irreducible representations). We will be calling such reducible representations that are obtained as limits of irreducible ones **liminal**.

This manner of reasoning is how one will be passing from station $\boxed{5}$ to station $\boxed{4}$.

4.6 Herbrand’s Theorem; regular primes, properly irregular primes, and *improperly irregular primes*

A consequence of the classical Herbrand Theorem is that if an odd prime number p does *not* divide any of the Bernoulli numbers B_{2k} for $k = 1, 2, \ldots, p - 3$ then p does not divide the class number of the cyclotomic field $\mathbf{Q}(e^{2\pi i/p})$. The bare bones of the argument for this is the following: first one constructs a specific element θ in the integral group ring of the Galois group $\mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q})$ that annihilates the p-primary component of the ideal class group of $\mathbf{Q}(e^{2\pi i/p})$ and then one shows that under the hypotheses above, this element θ is a unit in the integral group ring. See Chapter 6 of Larry Washington’s *Introduction to cyclotomic fields* [**?**].

A prime number p that does not divide the class number of the cyclotomic field $\mathbf{Q}(e^{2\pi i/p})$ is called a **regular prime**. It is known that there are infinitely many irregular primes but even now, after much thought has been devoted to this area of mathematics, it is unknown whether or not there are infinitely many regular primes. For the story of conjectures regarding this and conjectural ideas about the related statistics, see [**?**]. If a prime p is irregular, but does not divide the class number of the maximal totally real subfield, $\mathbf{Q}(e^{2\pi i/p} + e^{-2\pi i/p})$, of $\mathbf{Q}(e^{2\pi i/p})$, then p is called **properly irregular**. A conjecture of Vandiver ([**?**]; see also section 10.3 in [**?**]) is that *all* irregular primes are properly irregular. Vandiver’s conjecture has the advantage of being verified for $p < 12,000,000$ ([**?**]) and has strong consequences, so even if there exist *improperly irregular* primes, proper irregularity is a useful condition to bear in mind. What is of interest, specifically for our story, is that the existence of the Galois extensions constructed by Ribet was already known to be true for any properly irregular prime p.

5 Rounding the circuit with the prime 691

Here is a taste of the “six facets” discussed above, as they manifest themselves for the prime 691. Since 691 is properly irregular, as mentioned above, the bald existence of a Galois extension with the properties that Ribet’s theorem guarantees is not at issue: the desired Galois extension has been known to exist for quite a long time. Even more to the point—as we shall get to, later—for this very example, Serre had suggested exactly where to find the desired Galois extension, and unpublished work of Greenberg and Monsky did indeed find it, and in essentially the same context that we will be reviewing, all this happening before Ribet’s work.

Here, rather, we will have our sights on Ribet's general method of construction for all (irregular) primes p, but we concentrate on this well-studied $p = 691$. We will make the circuit and for each of the six vertices of our hexagon we will be signaling an explicit piece of number theory—a computation—related to that vertex. These six phenomena are—as we indicated in the introduction–essentially linked together.

$\boxed{1}$ Divisibility of the numerator of Bernoulli numbers modulo p.

Let B_k be the classical k-th Bernoulli number as in Jacob Bernoulli's famous *Ars Conjectandi*[11]: For odd integers $k > 1$, B_k vanishes. For even integers $2k$ the Bernoulli number B_{2k} is the coefficient of $x^{2k}/2k!$ in the power series expansion

$$\frac{x}{e^x - 1} \;=\; 1 - \frac{x}{2} + \sum_{k=1}^{\infty} (-1)^{k+1} B_{2k} \frac{x^{2k}}{2k!}.$$

We moderns have (as Bernoulli himself also had) easy methods for its computation, and here is a list of the first few $\frac{B_{2k}}{4k}$.

$B_2/4 \;=\; +1/24$

$B_4/8 \;=\; -1/240$

$B_6/12 \;=\; +1/504$

$B_8/16 \;=\; -1/480$

$B_{10}/20 \;=\; +1/264$

which might lead us to make a rash conjecture about the numerator of these numbers, if not for the next case:

$B_{12}/24 \;=\; -691/65520$

The prime $p = 691$*, then, divides the numerator of* $B_{12}/24$. Hold that thought!

$\boxed{2}$ The ideal class group taken modulo p of the cyclotomic field $\mathbf{Q}(e^{2\pi i/p})$.

For a number field K the set of nonzero ideals in its ring of integers, viewed with the multiplicative structure that it naturally has (multiplication of ideals)—but taken modulo the equivalence relation generated by saying that any principal ideal is trivial—forms a finite abelian group, called the *ideal class group* of K.

[11] An English translation of this extraordinary work has recently been published [?] with extensive introductory material.

Keep to $p = 691$ but in the body of this discussion we will revert to denoting 691 simply by the letter p. Take K to be the cyclotomic field $\mathbf{Q}(e^{2\pi i/p})$. The automorphism group of this field—that is, the Galois group $\mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q})$— is canonically isomorphic to $\mathbf{F}_p^*$, the multiplicative group of the prime field of characteristic p. Denote this canonical isomorphism by:

$$\iota : \mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q}) \longrightarrow \mathbf{F}_p^*.$$

Form $X :=$ the ideal class group of K modulo $p(= 691)$ which has a natural action of $\mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q})$ which we denoted $(\alpha, x) \mapsto \alpha(x)$ for $\alpha \in \mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q}) = \mathbf{F}_p^*$ and $x \in X$.

What is the abelian group X (which is convenient to think of as an $\mathbf{F}_p$-vector space) and what is this action?

The answer is that X is nontrivial[12], which already signals that the *elementary strategy* for proving Fermat's Last Theorem—the method used for the prime exponent 3 by Euler[13] —will not work for the exponent 691. Hold that thought!

The group X viewed as vector space over $\mathbf{F}_p$ is, in fact, of dimension two, and has a basis $\{x, y\}$ of eigenvectors for the action of

$$\mathrm{Gal}\big(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q}\big) \ \xrightarrow{\iota}\ \mathbf{F}_p^*$$

such that for $\alpha \in \mathrm{Gal}\big(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q}\big)$ we have the formulas

$$\begin{aligned} \alpha(x) &= \iota(\alpha)^{691-12} \cdot x = \iota(\alpha)^{-11} \cdot x \\ \alpha(y) &= \iota(\alpha)^{691-200} \cdot y = \iota(\alpha)^{-199} \cdot y \end{aligned}$$

Though both of these "lines" in the two-dimensional $\mathbf{F}_p$ vector space X are interesting, the discussion about each of them is somewhat similar, so for specificity we sometimes concentrate, below, a bit more on the line $x \cdot \mathbf{F}_p$ in X admitting an action by $\mathrm{Gal}\big(\mathbf{Q}(e^{2\pi i/p})\big)$ via the character ι^{-11}.

$\boxed{3}$ Abelian unramified extensions of order 691 over the cyclotomic field $K = \mathbf{Q}(e^{2\pi i/691})$

There are field extensions L/K that are (cyclic) of order 691, that are everywhere unramified and that have the further property that $L/\mathbf{Q}$ is Galois[14] and furthermore—there are precisely two

[12] Therefore it might be an amusing—but possibly a time-consuming and conceptually not very rewarding—task to directly seek specific nonprincipal ideals whose p-th powers are principal. This is—in effect—elegantly done for us, somehow, by "Herbrand-Ribet" both in our particular case ($p = 691$) and (the analogous task is done) in the general case of *irregular* primes p.

[13] Ernst Kummer already understood the deep arithmetic consequences that follows from knowledge of the *vanishing* of the ideal class group modulo p of the cyclotomic number field $\mathbf{Q}(e^{2\pi i/p})$. (If this happens, then "enough" of the fundamental theorem of arithmetic holds in the ring of integers of $\mathbf{Q}(e^{2\pi i/p})$ to allow one to prove Fermat's Last Theorem for the equation $x^p + y^p = z^p$.)

[14] Kummer knew that for p a prime number and for any field F of characteristic different from p that contained a primitive p-th root of unity, the cyclic field extensions of F of degree p are obtained by extracting p-th roots of elements of F; moreover, these cyclic field extensions are in one:one correspondence with the $\mathbf{F}_p$-lines in the $\mathbf{F}_p$–vector space $F^*/(F^*)^p$ (although he would express this knowledge via completely different vocabulary). So—for example—there are infinitely many cyclic extensions of degree 691 over the cyclotomic field $K = \mathbf{Q}(e^{2\pi i/691})$. One conceivable way, then, of pinpointing the *everywhere unramified* cyclic extensions of degree 691—there being exactly *two* such extensions that are Galois over $\mathbf{Q}$—would be to exhibit units υ in $K = \mathbf{Q}(e^{2\pi i/691})$ (υ not a perfect 691-st power) such that the field extensions obtained by extracting a 691-th root of υ are the ones you want. As we shall see, the method we are discussing does not exhibit these extensions in that way.

such field extensions L/K; call them $L^{\{12\}}$ and $L^{\{200\}}$. These two extensions are distinguished by the nature of the action by *conjugation*; that is, by the conjugation-action of any lifting $\tilde{\alpha}$ of $\alpha \in \mathrm{Gal}(K/\mathbf{Q})$ to $\mathrm{Gal}(L/\mathbf{Q})$ on elements $\gamma \in \mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/\mathbf{Q})$. This conjugation-action is given for $L = L^{\{12\}}$ by the formula:

$$\tilde{\alpha}\gamma\tilde{\alpha}^{-1} \; = \; \iota(\alpha)^{691-12} \cdot \gamma = \; \iota(\alpha)^{-11} \cdot \gamma,$$

and for $L = L^{\{200\}}$ it is given by:

$$\tilde{\alpha}\gamma\tilde{\alpha}^{-1} \; = \; \iota(\alpha)^{691-200} \cdot \gamma = \; \iota(\alpha)^{-199} \cdot \gamma.$$

$\boxed{4}$ Galois representations into $\mathrm{GL}_2(\mathbf{F}_{691})$

For each of these two extensions $L/\mathbf{Q}$ (i.e., $L = \; L^{\{12\}}$ or $L^{\{200\}}$) the equations above allow us to view $\mathrm{Gal}(L/\mathbf{Q})$ as a semi-direct product

$$\mathrm{Gal}(L/\mathbf{Q}) = \mathrm{Gal}(L/K) \ltimes \mathrm{Gal}(K/\mathbf{Q}).$$

From each of these extensions $L/\mathbf{Q}$ we now will describe corresponding two-dimensional representations of the Galois group of the rational field $\mathbf{Q}$, and—as we shall see—these "Galois representations" deserve our attention.

For example, take $L := L^{\{12\}}$. Let $\psi : \mathrm{Gal}(L/K) \simeq \mathbf{F}_p^+$ be a choice of isomorphism. Then form the representation

$$\bar{\rho} : \mathrm{Gal}(L/\mathbf{Q}) \; \longrightarrow \; \mathrm{GL}_2(\mathbf{F}_p)$$

by sending $\gamma \cdot \alpha$ to

$$\begin{pmatrix} 1 & \psi(\gamma) \\ 0 & \iota(\alpha)^{11} \end{pmatrix}$$

The displayed equations above show that this is indeed a homomorphism. Working similarly with L^{200} in this way, we get two Galois representations for our two choices of L; call them $\rho^{\{12\}}$ and $\rho^{\{200\}}$ respectively. The equivalence class[15] of each of these representations $\bar{\rho}$ is independent of our choices. Remember these representations!

$\boxed{5}$ The cuspform Δ of level 1 and weight 12.

I'm referring to that very intensely studied infinite product

$$\Delta(q) \; = \; q\prod_{n=1}^{\infty}(1-q^n)^{24} \; = \; 0 + \sum_{n=1}^{\infty}\tau(n)q^n$$

which can be directly thought of as it is presented here, namely, a power series in the variable q; or: putting $q = e^{2\pi i z}$ we may view it as an analytic function of the variable z in the upper half-plane,

[15] Recall that for any commutative ring with unit, R, and any group G, two homomorphisms $h_1, h_2 : G \; \rightarrow \mathrm{GL}_N(R)$ are said to define **equivalent representations** if there is an element $\gamma \in \mathrm{GL}_N(R)$ such that h_2 is the composition of h_1 with the automorphism of $\mathrm{GL}_N(R)$ given by conjugation by γ.

where it is shown to satisfy the symmetry $\Delta(1/z) = z^{12}\Delta(z)$ (i.e., Δ is a classical modular form of level 1 and weight 12). Its Fourier series coefficients, $n \mapsto \tau(n)$, were brilliantly analyzed by generations of mathematicians—including Ramanujan. Simple recurrence relations (first described by Mordell) allow you to "retrieve" the Ramanujan "tau"-function $n \mapsto \tau(n)$ from the values $\ell \mapsto \tau(\ell)$ for all prime numbers ℓ.

A classical congruence gives us that

$$\tau(n) \equiv \sum_{0<d \mid n} d^{11} \mod 691$$

for every positive integer n. In particular,

$$\tau(\ell) \equiv 1 + \ell^{11} \mod 691$$

for every prime number ℓ.

There is also a unique cuspform—call it $\Delta^{\{200\}}$—of level 1 and weight 200 with Fourier coefficients $n \mapsto \tau^{\{200\}}(n) \in \mathbf{Q}_{691}$ that enjoys a similar congruence[16]:

$$\tau^{\{200\}}(n) \equiv \sum_{0<d \mid n} d^{199} \mod 691$$

for every positive integer n. In particular,

$$\tau(\ell) \equiv 1 + \ell^{199} \mod 691$$

for every prime number ℓ.

Remember this!

$\boxed{\mathbf{6}}$ The Eisenstein series E_{12} of level 1 and weight 12, modulo $p = 691$.

The Fourier expansion of this modular form is given by

$$E_{12}(q) = -B_{12}/24 + \sum_{n=1}^{\infty}\Big(\sum_{0<d \mid n} d^{11}\Big)q^n.$$

Since, as we've mentioned in $\boxed{\mathbf{1}}$, (i.e., we are coming full circle) $B_{12}/24 = -691/65520$. So $B_{12}/24 \equiv 0$ modulo 691 and therefore our Eisenstein series has zero as its constant term modulo 691. Since, as we've mentioned in $\boxed{\mathbf{4}}$,

$$\tau(n) \equiv \sum_{0<d \mid n} d^{11} \mod 691$$

[16] I'm thankful to William Stein for computing this: he tells me that the Fourier coefficients $\{\tau^{\{200\}}(n)\}_n$ generate a field of degree 16 over $\mathbf{Q}$ and the primes dividing 691 in this field have degrees $1, 1, 2, 2$, and 10; completion at one of those primes of degree 1 gives us the cuspform $\Delta^{\{200\}}$ with Fourier coefficients in $\mathbf{Q}_{691}$ enjoying the congruences we are describing.

we get that the Eisenstein series E_{12} has *exactly* the same Fourier expansion, modulo 691, as Δ.

Similarly, the Eisenstein series E_{200} of level 1 and weight 200 given by

$$E_{200}(q) = -B_{200}/400 \; + \; \sum_{n=1}^{\infty} \Big(\sum_{0<d \; | \; n} d^{199} \Big) q^n$$

has the property that the numerator of $-B_{200}/400$, its constant term[17] is divisible by the prime 691. So, again, the Fourier series of the Eisenstein series E_{200} taken modulo 691 has zero as constant term[18].

The Herbrand-Ribet Theorem assures us that these six arithmetic events are not mere coincidences, and that there are implications of a very similar sort going all the way around the hexagon for any prime number p.

6 Rounding the circuit again: a statement of the Herbrand-Ribet Theorem

Theorem 1. *(Herbrand-Ribet) Let p be a prime number and $2k$ an even integer greater than 2 and less than $p-1$. These are equivalent:*

$\boxed{\mathbf{1}}$ *The numerator of the "Bernoulli number" $\frac{B_{2k}}{4k}$ is divisible by p. In Kummer's terminology, this is what is meant by p being an* **irregular** *prime number*[19].

$\boxed{\mathbf{6}}$ *The constant term in the Fourier expansion of the Eisenstein series E_{2k} of weight $2k$ and level 1 is congruent to zero modulo p. (Colloquially we can say: the Fourier expansion of E_{2k} "looks cuspidal modulo p.")*

$\boxed{\mathbf{5}}$ *For some—or, equivalently, for every—integer $2 \leq w < \infty$ there exist the following objects:*

[17] which, if you're curious, is: $389 \cdot 691 \cdot 5370056528687$ times this 204-digit prime number: 3452690329392158031464109281736967404068448156842396721012992064214519445919256941544565276067662360108 ∼ 74972724155570842527652727868776362959519620872735612200601036506871681124610986596878180738901486527

[18] In fact, E_{200} has *exactly* the same Fourier expansion, modulo 691, as a certain cuspform of weight 200—that we will call $\Delta^{\{200\}}$— has. This is meant in the sense that the Fourier coefficients of $\Delta^{\{200\}}$ lie in the ring of integers of a number field that has a degree one prime ideal whose residue field is $\mathcal{F}_{691}$; when reduced modulo this prime ideal one has $\Delta^{\{200\}} \equiv E_{200}$.

[19] therefore—as someone mentioned after my lecture—-in contrast to what happens in elementary geometry, *our* hexagon is only of special interest when it is *irregular*... I apologize for not crediting this person: I've forgotten who it was, but wouldn't be surprised if he or she may prefer to remain anonymous.

- *a number field $F_w \subset \bar{Q}$. Let $\mathcal{O}_w \subset F_w$ denote the ring of integers in F_w.*
- *A prime ideal $P_w \subset \mathcal{O}_w$ such that $\mathcal{O}_w/P_w = \mathbf{F}_p$.*
- *A power series*

$$\Phi_w \;=\; 0 + q + \sum_{n=2}^{\infty} t_w(n) q^n$$

  with coefficients $t_w(n) \in \mathcal{O}_w$, such that

  - *when viewed as a power series in $q = e^{2\pi i z}$ with coefficients in $\mathbf{C}$ via any imbedding $F_w \hookrightarrow \mathbf{C}$ this power series is the Fourier series of a cuspidal eigenform on $\Gamma_1(p)$ of weight w;*
  - *when reduced modulo the prime ideal P_w and viewed as a power series with coefficients in the prime field $\mathbf{F}_p$ we have, for each integer $n \geq 0$, the congruence:*

$$t_w(n) \equiv \sum_{0<d \ | \ n} d^{2k-1}.$$

    In more colloquial vocabulary, the Fourier expansion of Φ_w is congruent to the Fourier expansion of E_{2k} modulo p.

$\boxed{\mathbf{4}}$ *There exist the following objects:*

- *a number field $F \subset \bar{Q}$. Let $\mathcal{O} \subset F$ denote the ring of integers in F.*
- *A prime ideal $P \subset \mathcal{O}$ such that $\mathcal{O}/P = \mathbf{F}_p$. Let $\mathcal{O}_P \subset F_P$ be the completions of $\mathcal{O} \subset F$ at P. Let $\pi_P \in \mathcal{O}_P$ denote a uniformizer of the discrete valuation ring $\mathcal{O}_P$.*
- *An abelian variety A defined over $\mathbf{Q}$ with the following properties:*
  - *The ring of integers $\mathcal{O}$ acts as a ring of endomorphisms of A over $\mathbf{Q}$.*
  - *Letting $A[P^\nu] \subset A(\bar{Q})$ denote the intersection of the kernels of all the endomorphisms of A that lie in the ideal P^ν and noting that $\pi_P : A[P^{\nu+1}] \to A[P^\nu]$ is a surjective homomorphism, consider the projective system*

$$\cdots \to A[P^{\nu+1}] \to A[P^\nu] \to \cdots A[P].$$

    The projective limit $T_P(A) := \lim_\nu A[P^\nu]$ has a natural $\mathcal{O}_P[G_{\mathbf{Q}}]$-action. Form the F_P-vector space $V := V_P(A) = T_P(A) \otimes_{\mathcal{O}_P} F_P$. This F_P vector space is of dimension two.
  - *There is an $\mathcal{O}$-lattice $\Omega \subset V$ stable under the action of $G_{\mathbf{Q}}$ for which the representation*

$$G_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(\Omega/\pi_P\Omega) \cong \mathrm{GL}_2(\mathbf{F}_p)$$

    is indecomposable but reducible and whose semisimplification consists of two characters

$$\mathbf{1}, \omega_p^{2k-1} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_1(\mathbf{F}_p) = \mathbf{F}_p^*$$

    *where $\mathbf{1}$ is the trivial character and ω_p is the basic "p-cyclotomic" character introduced in subsection **??** above.*

- *A has good reduction at all primes different from p.*
- *A achieves good reduction at all primes when the base field is extended to the maximal real subfield* $\mathbf{Q}(e^{2\pi i/p}+e^{-2\pi i/p})$ *in the p-cyclotomic field.*

$\boxed{\mathbf{3}}$ *Let K be the cyclotomic field* $K=\mathbf{Q}(e^{2\pi i/p})$. *There is a field extension* L/K *that is cyclic of order p, that is everywhere unramified, and that has the further property that* $L/\mathbf{Q}$ *is Galois, and furthermore the action by conjugation of any lifting* $\tilde{\alpha}$ *of* $\alpha\in\mathrm{Gal}(K/\mathbf{Q})$ *to* $\mathrm{Gal}(L/\mathbf{Q})$ *on any element* $y\in\mathrm{Gal}(L/K)\subset\mathrm{Gal}(L/\mathbf{Q})$ *is given by the formula:*

$$\tilde{\alpha}y\tilde{\alpha}^{-1} \;=\; \iota(\alpha)^{p-2k}\cdot y = \;\iota(\alpha)^{1-2k}\cdot y.$$

$\boxed{\mathbf{2}}$ *There is a cyclic subgroup X of order p in the ideal class group of* $K=\mathbf{Q}(e^{2\pi i/p})$ *stabilized by the action of the Galois group* $\iota:\mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q})\simeq\mathbf{F}_p^*$ *and such that the action of* $\alpha\in\mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q})$ *on X is given by the formula:*

$$\alpha(x) \;=\; \iota(\alpha)^{p-2k}\cdot x \;=\; \iota(\alpha)^{2k-1}\cdot x.$$

We end part I of this article by accompanying these statements in their counterclockwise route with some brief discussion; a few of these points will be gone into in slightly more detail in the later sections.

- $\boxed{\mathbf{1}}\Rightarrow\boxed{\mathbf{6}}$.

  This is just because $\frac{B_{2k}}{4k}$ *is* the constant term of the Fourier expansion of E_{2k}.

- $\boxed{\mathbf{6}}\Rightarrow\boxed{\mathbf{5}}$.

  Here Ken depends upon the algebraic-geometric interpretation of modular forms—as sections of bundles over algebraic curves. The question of *lifting* a mod p form to characteristic zero is then dealt with by standard exact sequences, very much helped by the fact that we are over *algebraic curves*. (The analogous procedure when one deals with higher rank automorphic forms is not so smooth-going.) Once one lifts, standard methods allow one to obtain a lift is an eigenform. i.e., is "eigen" for all the Hecke operators that one needs.

  A curious point is that we have our choice of infinitely many different newforms f (at least one for each weight). Ken chose to work with cuspidal newforms of weight two. In the next step we will be making use of Galois representations associated to these cuspidal eigenforms, and in the case of weight two, these representation are somewhat concretely obtained as a representations occurring in the natural action of Galois on the p-power torsion points of abelian varieties.

  In particular, in our discussion of $p=691$, we chose to single out Δ, the eigenform of weight twelve (that we will also call Φ_{12} in section **??** below, but Ken worked with weight two—and in that specific instance he would be visualizing the Galois representation obtained by the natural Galois action on the 691-power torsion points of the abelian variety of dimension 2508 that we called A, attached to the eigenform we will be calling Φ_2 in section **??**.

- $\boxed{\mathbf{5}} \Rightarrow \boxed{\mathbf{4}}$.

  Let $f = \sum_{n=1}^{\infty} a_n(f)q^n$ be a cuspidal newform of level p (i.e., on $\Gamma_1(p)$) of weight > 1. The theorem of Deligne (see section **??** below) provides us with degree two Galois representations as discussed previously in the context of Δ. In particular—if we view the Fourier coefficients of f as lying in some finite discrete valuation ring extension D of $\mathbf{Z}_p$ (this generally requires making a choice) then Deligne will give us a representation $\rho_D : G_{\mathbf{Q}} \to \mathrm{GL}_2(D)$ which is

  - irreducible as a representation over the field of fractions of D, and
  - which has the property that the trace of Frobenius at ℓ (for primes $\ell \neq p$) is equal to $a_\ell(f)$.

  If f is the eigenform given to us by $\boxed{\mathbf{4}}$ then

  - the residue field of D is $\mathbf{F}_p$,
  - the Fourier expansion of f modulo the maximal ideal of D is congruent to the Fourier expansion of E_{2k} modulo p making
  - the Galois representation $\bar{\rho}_D$ (i.e., ρ_D viewed over $\mathbf{F}_p$) *reducible*; with, in fact, item its semisimplification $\bar{\rho}_D^{\mathrm{ss}}$ equal to the trivial character 1 plus ω^{2k-1} where ω is the *cyclotomic character modulo* p, i.e., ω is the composition:

  $$G_{\mathbf{Q}} \to \mathrm{Gal}(\mathbf{Q}(e^{2\pi i/p})/\mathbf{Q}) \overset{\iota}{\simeq} \mathbf{F}_p^*.$$

- $\boxed{\mathbf{4}} \Rightarrow \boxed{\mathbf{3}}$.

  So far, this seems like a very unpromising picture, for we have no idea whether the representation $\bar{\rho}_D$ is nothing more than the sum $\mathbf{1} \oplus \omega^{2k-1}$, which would be a nice-enough fact, but would not give us what we want; namely a *nontrivial* abelian Galois extension of the p-cyclotomic field (and, in fact, we even want more: we want it to be everywhere unramified and acted upon by Galois in a very particular way).

  It is at this point that the Ribet wrench comes to help us. Ken uses it to show that since the representation over F, the field of fractions of D, is *irreducible* he can *change* the D-lattice of the underlying F-vector space of the representation $\rho_D \otimes F$ to obtain a new $G_{\mathbf{Q}}$-stable D-lattice such that the corresponding $\mathbf{F}_p$-representation obtained by reduction modulo the maximal ideal is *indecomposable* and in fact—in this situation—one can do this in two ways: either so as to have the trivial representation $\mathbf{1}$ as *sub*-representation of the corresponding indecomposable residual representation or as *quotient*. Ken goes with *quotient* and uses the algebraic geometry of the ‘ to guarantee that he has constructed precisely the type of abelian extension that is required by $\boxed{\mathbf{3}}$. We call this the **preferred indecomposable residual representation**. We shall be going into a few details about this last point in section **??** below.

- $\boxed{\mathbf{3}} \Rightarrow \boxed{\mathbf{2}}$.

  This leg of the journey is guaranteed by Class Field Theory as discussed briefly in subsection **??** above.

- $\boxed{\mathbf{2}} \Rightarrow \boxed{\mathbf{1}}$.

  This is Herbrand's theorem (proved before World War II).

I hope that my article has, up to this point, conveyed to readers of general background the flavor of some ideas behind this type of explicit construction of abelian extensions. The remaining sections of this article will go a bit further into some of the techniques required for–and connected to—the proof, and also current related work.

Part II: Galois representations coming from Algebraic Geometry

7 Galois extensions of number fields, prime splitting phenomena, and Frobenius conjugacy classes

If L/K is a finite Galois extension of number fields, with $G = \mathrm{Gal}(L/K)$ the action of the group G on L stabilizes the ring of integers $\mathcal{O}_L$ in L and fixes the subring $\mathcal{O}_K = \mathcal{O}_L \cap K$ of integers in K. Of major arithmetic interest is the manner in which primes P of K (i.e., nonzero prime ideals of $\mathcal{O}_K$) split (or not) in L. A historically important case of this general problem is answered by the theorem ascribed to Fermat that says that a prime number is expressible as a sum of two squares (i.e., splits into a product of two primes in the field of gaussian numbers) if and only if it is not congruent to -1 mod 4. In general, the splitting of P in the field L is given by the prime factorization of the $\mathcal{O}_L$-ideal $P \cdot \mathcal{O}_L$. Such a factorization will look like:

$$P \cdot \mathcal{O}_L \;=\; Q_1^e \cdot Q_2^e \cdot \dots \cdot Q_m^e$$

where the Q_i are mutually distinct primes of L and $e \geq 1$.

One usually refers to the primes Q_i that occur in this formula as the "primes Q of L *lying above* P."

For any prime P of K the action of the Galois group G on $\mathcal{O}_L$ induces a *transitive* action on the set of primes of L lying above P. For any of these Q's, $G_Q \subset G$ denotes the isotropy subgroup at Q of the action —i.e., $G_Q := \{g \in G \mid g \cdot Q = Q\}$. Fixing, then, a prime Q lying above P we may identify the set of all primes of L lying above P with the left coset space G/G_Q in the evident way. The isotropy subgroup G_Q stabilizes both $\mathcal{O}_L$ and $Q \subset \mathcal{O}_L$ and therefore one induces a natural action of G_Q on the finite ("residue field at Q") $\kappa_Q := \mathcal{O}_L/Q$.

The prime P is said to be **ramified** in L/K if $e > 1$; there are only finitely many primes P ramified in L/K (these being the primes dividing the discriminant $\mathrm{Disc}(L/K)$ which is a nonzero ideal of $\mathcal{O}_K$. If, for any nonzero ideal $I \in \mathcal{O}_K$ we denote by NI its norm (meaning the cardinality of the set $\mathcal{O}_K/P$) then a sufficient condition for P to be unramified is that the integer NP not divide $N\mathrm{Disc}(L/K) \neq 0$.

In the case where P is unramified in L/K the isotropy group G_Q for any Q lying above P has a strikingly precise structure:

> The isotropy group G_Q is *cyclic* with a canonical generator (called the **Frobenius element at Q**
>
> $$Frob_Q \;\in\; G_Q \;\subset\; G$$
>
> determined uniquely by the property that the natural action of $Frob_Q$ on the residue field at Q, $\mathcal{O}_L/Q$, consists in raising every element to the NP-th power; equivalently:
>
> $$Frob_Q(x) \;\equiv\; x^{NP} \quad \bmod Q \quad \text{for } x \;\in \mathcal{O}_L.$$

The set of elements $Frob_Q \in G$ where Q ranges through all primes of L lying above P consists in a single conjugacy class of G—determined by P— and we will refer to this conjugacy class $c_P = c_P(L/K) \subset \mathrm{Gal}(L/K)$ as the **Frobenius conjugacy class (relative to the Galois extension L/K)** attached to the (unramified) prime P of K. This assignment is nicely functorial for nested Galois extensions of K; i.e., if $K \subset L \subset M$ are fields with M/K and L/K Galois, we have—for all primes P of K unramified in M/K—that $c_P(L/K)$ is the image of $c_P(M/K)$ under the natural surjection $\mathrm{Gal}(M/K) \to \mathrm{Gal}(L/K)$.

The Galois group $G = \mathrm{Gal}(L/K)$ of a Galois extension L/K of number fields comes—then—with an impressive amount of extra structure. There are many ways of bottling this extra structure but here is a way that I find helpful.

By a **Cebotarev Group** over a number field K I will mean a finite group G together with a function $P \mapsto c_P$ defined on almost all (i.e. all but a finite number of) primes P over K that associates to P a conjugacy class c_P in G, and that has the property that for any conjugacy class $c \subset G$, there are infinitely many P with $c_P = c$ and more precisely:

$$\lim_{X\to\infty} \frac{|\{P \mid c_P = c \text{ and } NP < X\}|}{|\{P \mid NP < X\}|} \;=\; \frac{|c|}{|G|}$$

where the absolute value sign around the symbol for a set means its cardinality.

Let us say that two Cebotarev groups over K, $\{G; P \mapsto c_P\}$ and $\{G'; P \mapsto c'_P\}$, are isomorphic if there is an isomorphism $G \simeq G'$ which sends c_P to c'_P for all but finitely many primes P for which both c_P and c'_P are defined.

We can formulate a version of the famous *Theorem of Cebotarev* as follows:

Theorem 2. *Let L/K be a Galois extension of number fields with $G = \mathrm{Gal}(L/K)$. The rule that assigns to each prime P of K that is unramified in L/K the Frobenius conjugacy class $c_P = c_P(L/K)$ defines a Cebotarev group $\{G; P \mapsto c_P\}$. The isomorphism class of this Cebotarev group determines L/K up to isomorphism.*

The first sentence of the theorem is the classical Cebotarev Theorem[20]. The second assertion is seen as follows. Let $\bar{K}/K$ be an algebraic closure of K.

[20] For a proof of the theorem cf. Thm. 10 Ch. VIII, section 4 of [**?**]; also see section 2.7 of [**?**] which casts the theorem in the context of global fields. For Cebotarev's original article see [**?**]; and to get a sense of his generosity of spirit see the extraordinary tale beginning on page 131 in [**?**].

Lemma 1. *Let $L, L' \subset \bar{K}$ be subfields each of which is a finite Galois extension field of K and suppose that there is an isomorphism $\psi : \mathrm{Gal}(L/K) \simeq \mathrm{Gal}(L'/K)$ such that if P runs through almost all primes of K which are unramified in L/K and in L'/K the isomorphism ψ sends the Frobenius conjugacy class relative to the Galois extension L/K, $c_P \subset G = Gal(L/K)$, to the Frobenius conjugacy class relative to the Galois extension L'/K, $c'_P \subset G' = \mathrm{Gal}(L'/K)$. Then $L' = L$.*

Proof: Let $M = L \cdot L' \subset \bar{K}$ be the compositum. Put $G_M := \mathrm{Gal}(M/K)$ and form

$$G_M \hookrightarrow G \times G',$$

the injection being the product of the natural surjections onto each factor.

Suppose first that the natural surjection $G_M \to G$ is not an isomorphism. Let $\mathcal{C} \subset G_M$ denote the conjugacy class in G_M containing some nontrivial element of $N := \ker\{G_M \to G\}$. The natural projection $N \to G'$ is injective since $G_M \hookrightarrow G \times G'$ is injective. Let $c' \subset G'$ be the conjugacy class of G' containing the image of $\mathcal{C}$ noting that c' is not the (conjugacy class of the) identity in G'. By the classical Cebotarev Theorem for M/K every conjugacy class of G_M is the Frobenius conjugacy class of infinitely many primes of K. Therefore, by the functoriality of Frobenius conjugacy classes mentioned above, and by the hypotheses of our lemma, every conjugacy class of G_M maps to a conjugacy class of $G \times G'$ of the form (c, c') where $\psi(c) = c'$. But the conjugacy class $\mathcal{C}$ constructed above maps to (c, c') where c is the identity conjugacy class in G and c' is not the identity class in G'.

It follows that N is the trivial group, and therefore the natural projection $G_M \to G$ is an isomorphism, establishing our lemma.

It is traditional to study the structure packaged in the Cebotarev group associated to an extension L/K by considering linear representations of the underlying Galois group—i.e., choosing a Galois representation over K that is split by L/K—and using the vocabulary and techniques of analytic number theory. In fact, it is by such a route that the Cebotarev Theorem quoted above is proved.

Here is a brief hint of what is involved. Fix $\{G; P \to c_P\}$, a Cebotarev group over a number field K. Consider an (irreducible) complex representation $\eta : G \to \mathrm{GL}_n(\mathbf{C})$. For a prime number P for which c_P is defined, let

$$\mathcal{L}_{\eta,P}(T) := \det\bigl(1 - T \cdot M\bigr) \in \mathbf{C}[T])$$

be the characteristic polynomial of the $n \times n$ matrix $1 - T \cdot M \in \mathrm{Mat}_{n\times n}\bigl(\mathbf{C}[T]\bigr)$ where $M \in \mathrm{GL}_n(\mathbf{C})$ is any element of $\eta(c_P) \subset \mathrm{GL}_n(\mathbf{C})$. Thus,

$$\mathcal{L}_{\eta,P}(T) \;\equiv\; 1 - \mathrm{Trace}\bigl(\eta(Frob_Q)\bigr) \cdot T \quad \text{modulo higher powers of } T$$

where Q is any prime lying above P. Since η is a complex representation of a finite group, and since the Cebotarev theorem guarantees that every conjugacy class of G is c_P for (infinitely many) primes P, the representation η is already pinned down, up to isomorphism, by the data $P \mapsto \mathrm{Trace}\bigl(\eta(Frob_Q)\bigr)$ and therefore all the more by

$$P \mapsto \mathcal{L}_{\eta,P}(T).$$

From the above data, define a Dirichlet series (in the complex variable s)

$$\mathcal{L}(\eta, s) = \prod_P \mathcal{L}_{\eta,P}(NP^{-s})^{-1}.$$

Since the eigenvalues of any of the matrices M in the previous paragraph are roots of unity, a straightforward computation gives that the Dirichlet series $\mathcal{L}(\eta, s)$ converges in some right half-plane. Traditionally, one tries to learn detailed statistical information about the rule $P \mapsto c_P$ by establishing *further* analytic properties (e.g., analytic continuation, functional equation, location of poles and zeroes) of this collection of Dirichlet series

$$\eta \mapsto \mathcal{L}(\eta, s).$$

If the Cebotarev group in question comes from a Galois extension L/K we may view η as a representation of $\mathrm{Gal}(L/K)$ and $\mathcal{L}(\eta, s)$ is—except for factors corresponding to the missing primes P—the classical Artin L-function about which much is known, and even more is conjectured.

8 Towers of Galois representations; p-adic Galois representations

In the previous section we discussed *finite degree* Galois extensions of number fields, and corresponding Galois representations of finite Galois groups into $\mathrm{GL}_n(\mathbf{C})$. The focus, though, of much recent work is towards infinite degree extensions. For this we should say a few words about infinite Galois groups.

For K a field, choose an algebraic closure $\bar{K}$ of K and set $G_K := \mathrm{Gal}(\bar{K}/K)$ which we view as profinite topological group with its Krull topology; this is the topology for which closed subgroups of G_K are in one:one correspondence—as they would be in Galois theory of finite degree extensions—with the intermediate subfields of $\bar{K}/K$; any such closed subgroup $H \subset G_K$ *corresponds* to the subfield consisting of the elements of $\bar{K}$ fixed by all the elements of H. Since G_K is a profinite group, its continuous representations to Lie groups over $\mathbf{R}$ or over $\mathbf{C}$ necessarily factor through finite quotient groups; but this is no longer true if the target groups are, for example, Lie groups over p-adic fields. By a **p-adic Galois representation of degree n over K** we will mean a continuous representation

$$G_K \longrightarrow \mathrm{GL}_n(F)$$

where F is some extension field of finite degree over $\mathbf{Q}_p$. Since G_K is compact, one can show that any such continuous homomorphism can be conjugated to one that factors through $\mathrm{GL}_n(\mathcal{O}_F) \subset \mathrm{GL}_n(F)$ where $\mathcal{O}_F$ is the ring of integers (i.e., elements integral over $\mathbf{Z}_p$) in F. That is, in any equivalence class of such representations, there will be at least one homomorphism that has image in $\mathrm{GL}_n(\mathcal{O}_F)$. For example, if $F = \mathbf{Q}_p$, such a representation factors through a homomorphism

$$G_K \longrightarrow \mathrm{GL}_n(\mathbf{Z}_p)$$

which itself can be viewed as a projective limit of homomorphisms,

$$G_K \longrightarrow \mathrm{GL}_n(\mathbf{Z}/p^\nu\mathbf{Z}),$$

for $\nu = 1, 2, 3, \dots$, these representations being split by finite Galois extensions of K,

$$K \subset L_1 \subset L_2 \subset \cdots \subset L_\nu \subset \dots,$$

and if $L_\infty := \cup_\nu L_\nu$ then L_∞/K is a Galois extension of K whose—possibly infinite—Galois group is the compact p-adic Lie subgroup of $GL_n(\mathbf{Z}_p)$ that is the image of G_K in $\mathrm{GL}_n(\mathbf{Z}_p)$.

Often we will be content to deal with algebraic extension fields of $\mathbf{Q}_p$, i.e. subfields F of $\bar{\mathbf{Q}}_p$, an algebraic closure of $\mathbf{Q}_p$, but sometimes it is useful to allow a certain larger field as field of scalars, namely the field $\mathbf{C}_p := \hat{\bar{\mathbf{Q}}}_p$ the hat $\hat{}$ signifying the completion of $\bar{\mathbf{Q}}_p$ with respect to its p-adic valuation[21].

Usually we will be dealing with Galois representations that are *unramified except possibly at a finite number of places.* For such a Galois representation $\rho : G_K \longrightarrow \mathrm{GL}_n(F)$ we have a convenient "numerical handle" that determines ρ, up to semisimplification; namely, the function that associates to each place v of K unramified in ρ the value $a_\rho(v) := Trace_F(\rho(Frob_v)) \in F$. This function

$$v \;\mapsto\; a_\rho(v)$$

plays a central role in any dealings with a Galois representation ρ.

In particular, if $K = \mathbf{Q}$, we may view this function as taking values on "almost all" prime numbers ℓ, i.e., $\ell \;\mapsto\; a_\rho(\ell)$ and note that we've already had hints of such functions, such as the Ramanujan function $\ell \mapsto \tau(\ell)$, alluded to in our discussion of $\boxed{\mathbf{4}}$ above. An excellent general introduction to ℓ-adic representations is given in Serre's article [**?**] as well as in his earlier treatise [**?**].

9 Deligne's Theorem for the modular form Δ

A theorem of Deligne—a special case of which we shall be quoting below— gives us that for every prime number p and every modular eigenform, there is a continuous irreducible degree two p-adic Galois representation that is closely related to the eigenform in the sense that the Fourier coefficients of the eigenform determine—in a fairly direct way—the equivalence class of the representation. The previous two sections give us the vocabulary we need to discuss such connection between modular forms like

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

and the p-adic Galois representations that connect to them.

In this section—to focus ideas—we will concentrate on this classical form Δ itself, and in the next section we will specialize even further by considering $p = 691$. Our modular form Δ is related by the mod 691 congruence to the Eisenstein series E_{12} as we discussed in Part I. We will be introducing

[21] This field $\mathbf{C}_p$ sometimes is called "Tate's p-adic complex numbers" the *Tate* part of its name because Tate first defined and used it, the *complex numbers* part of its name because $\mathbf{C}_p$ is–deprived of its topology—abstractly isomorphic to the classical field of complex numbers; $\mathbf{C}_p$ is–in particular—algebraically closed, and—being a completion—is complete.

a representation denoted $\rho_{\Delta,691}$ that will be the key for us, in constructing the abelian field $L^{\{12\}}$ and the Galois representation $\rho^{\{12\}}$ that cuts it out. A very similar discussion beginning with the 691-adic cuspform $\Delta^{\{200\}}$ would construct the abelian field $L^{\{200\}}$ and the Galois representation $\rho^{\{200\}}$ that cuts it out (see footnote (18) below).

Essential input here is the famous theorem of Deligne relating modular eigenforms to Galois representations:

Theorem 3. *(Deligne) Let p be a prime number. There is a continuous irreducible degree two Galois representation*

$$\rho_{\Delta,p} : G_{\mathbf{Q}} \to \mathrm{Aut}_{\mathbf{Q}_p}(V) \approx \mathrm{GL}_2(\mathbf{Q}_p)$$

(where V is a 2-dimensional $\mathbf{Q}_p$-vector space) such that for all primes $\ell \neq p$ the representation $\rho_{\Delta,p}$ is unramified at ℓ and has the property that the trace of Frobenius at ℓ is equal to $\tau(\ell) \in \mathbf{Z} \subset \mathbf{Q}_p$.

Proof: See ([**?**]).

This condition (that the trace of Frobenius at ℓ is equal to $\tau(\ell) \in \mathbf{Z} \subset \mathbf{Q}_p$ for all primes $\ell \neq p$) determines the character of $\rho_{\Delta,p}$ since the Frobenius elements are dense, and therefore—since $\rho_{\Delta,p}$ is irreducible—the representation is pinned down if we know the Fourier coefficients of Δ. In the discussion to follow, first let us suppose that p is any (odd) prime number.

Suppose you are given a $G_{\mathbf{Q}}$-stable lattice $M \subset V$ (so M is a free $\mathbf{Z}_p$-module of rank two). By passing to $\bar{V} = M/pM = M \otimes_{\mathbf{Z}_p} \mathbf{F}_p$ we get an $\mathbf{F}_p$-representation of $G_{\mathbf{Q}}$ of degree two,

$$\bar{\rho}_{\Delta,p,M} : G_{\mathbf{Q}} \to \mathrm{Aut}_{\mathbf{F}_p}(\bar{V}) \approx \mathrm{GL}_2(\mathbf{F}_p)$$

that may—and in the cases of specific interest to us, *will*—depend upon the choice of lattice M. We will refer to $\bar{\rho}_{\Delta,p,M}$ as the **residual representation** obtained from $\rho_{\Delta,p}$ via the lattice M. Since—by the Cebotarev Density Theorem—the conjugacy classes of Frobenius elements (associated to all prime numbers $\ell \neq p$) run through *all* conjugacy classes of the Galois group of the splitting field of $\bar{\rho}_{\Delta,p,M}$; and since p is odd, the character of the (degree two) representation mod p, $\bar{\rho}_{\Delta,p,M}$, determines its semisimplification, we get that this semisimplification is completely characterized[22] by the function

$$\ell \ \mapsto \ \tau(\ell) \qquad \mathrm{mod}\ p.$$

For general prime numbers p the beauty of the representation $\rho_{\Delta,p}$ is that it is obtained naturally, and systematically, from the evident action of Galois on a piece of the étale cohomology group of an algebraic variety. The mild difficulty one encounters in working with this is that we are dealing with an H^{11}, i.e., cohomology in dimension eleven.

10 Moving from one weight to another

If you are particularly interested in the Galois representation modulo p, an idea, initially due to Serre, and also, in a slightly different formulation, to Koike, is to *replace* high weight modular

[22] See [**?**] and the proof—especially page 216—of (30.16), i.e., the Brauer-Nesbitt Theorem.

eigenforms by weight two eigenforms that have Fourier series that are congruent modulo p. One can do this! For background related to this, see Serre's paper ([**?**]) where he formulates his conjecture about Galois representations modulo p and modular forms mod p.

Now if you work with a weight two eigenform you may expect to find its associated Galois representation realized as the Galois module of one-dimensional cohomology with coefficients in $\mathbf{Q}_p$ of some simple abelian variety; or, to put it even more concretely, realized in the action of Galois on the p-torsion points of such an abelian variety[23]. Since the weight two modular form has Fourier coefficients congruent modulo p to the high weight eigenform you started out with, the mod p Galois representations associated to these modular forms will be equivalent, at least after semisimplification, as discussed in Part I (**??**).

Moving in this way to weight two modular eigenforms and their associated abelian varieties you have, at least, the sensation that you are dealing with more concrete entities (than higher dimensional varieties and the Galois representations on their étale cohomology groups) and indeed: in certain instances working with modular forms of weight two and their associated abelian varieties actually does confer the advantage of significantly more control.

11 Returning to $p = 691$

Here is what happens in the special case $p = 691$: since as we have mentioned,

$$\tau(\ell) \equiv 1 + \ell^{11} \mod 691$$

for every prime number ℓ, we get that the semisimplification of $\bar{\rho}_{\Delta,691,M}$ is equivalent to the semisimplification of the representation $\bar{\rho}$ of $\boxed{\mathbf{4}}$; namely, to $\mathbf{1} \oplus \omega^{11}$. This is true for the representations $\bar{V} = M/pM$ obtained from *every* $G_{\mathbf{Q}}$-stable lattice $M \subset V$. In the 1967/1968 Séminaire Delange-Pisot-Poitou, Jean-Pierre Serre wrote that it seemed probable to him that the image of inertia at p under the representations $\bar{\rho}_{\Delta,691,M}$ was distinct from the image of the full Galois group, thereby generating an everywhere unramified cyclic extensions of $\mathbf{Q}(e^{2\pi i/691})$ of degree 691. Serre went on to suggest that, perhaps, one could determine this using an analysis of the representation attached to Δ modulo 691^2 (page 507 of [**?**]). This was the approach similar to one taken up later by Ralph Greenberg and Paul Monsky who indeed constructed the sought-for everywhere unramified cyclic extension[24].

[23] There will, in general, be more than one such simple abelian variety, but (at least) one of them will be a quotient of $J_1(p)/J_0(p)$; here we have used standard notation.

[24] This is unpublished. Using the fact that $p = 691$ is properly irregular (and therefore the converse to Herbrand's Theorem is known for $p = 691$) and also using facts about the representation attached to Δ modulo 691^2 which guarantees that up to homothety that are *only two* Galois stable lattices M as above, and and the action of the Galois group $G_{\mathbf{Q}}$ on M/pM for each of these lattices is triangular and non-semisimple (with diagonal characters 1 and ω^{11} occurring in the two different orders in the two lattices) Greenberg and Monsky showed that the piece of the 691-Hilbert class field of $\mathbf{Q}(e^{2\pi i/691})$ corresponding to the character ω^{-11} is contained in the field extension of $\mathbf{Q}$ that we called L^{12} in part I) cut out by Deligne's 691-adic representation of $G_{\mathbf{Q}}$ associated with Δ. Greenberg wrote to me: "I was quite excited by the idea behind this at the time because I thought that it suggested a very promising approach to the converse."

12 Moving from Δ to eigenforms of weight $2, 3, 4, \ldots$ (when $p = 691$)

In the previous paragraph we were considering any prime number p and any eigenform. But if you fix attention to the weight 12 eigenform Δ and $p = 691$ you are in a quite a nice situation. Here, thanks to Hida's theory (references for which are [?], [?], [?]) there will be, for any weight $w > 1$ a 691-adic cuspidal eigenform Φ_w of weight w on $\Gamma_1(p)$ with Fourier coefficients in $\mathbf{Z}_p$ and whose Fourier expansion is of the form

$$1 \cdot q^1 + \sum_{n=2}^{\infty} a_w(n) q^n$$

where the cofficients a_w are *algebraic numbers* in $\mathbf{Q}_{691}$, and this Fourier series is congruent modulo 691 to the series

$$0 \ + \ \sum_{n+1}^{\infty} (\sum_{0<d \ | \ n} d^{11}) q^n$$

modulo 691, i.e., Φ_w has the same Fourier expansion modulo 691 as E_{12} or Δ. (Indeed, $\Phi_{12} = \Delta$).

If Φ_w is the 691-adic modular form of weight w for $w = 2, 3, 4, \ldots$ as above, denote by $\mathcal{F}_w \subset \mathbf{Q}_{691}$ the smallest field in $\mathbf{Q}_{691}$ generated (over $\mathbf{Q}$) by all the Fourier coefficients, $\{a_w(n); \ n = 1, 2, \ldots\}$, of Φ_w. It is known that $\mathcal{F}_w$ is a "number field," i.e., is of finite degree over $\mathbf{Q}$.

The above discussion is given (and in much greater generality) by *Hida Theory*.

Furthermore, in our case, the 691-adic cuspidal form Φ_w with the properties listed above is unique, for each $w > 1$; the appendix below sketches a proof of this. Deligne's theorem provides us with an infinite sequence of 691-adic representations,

$$\rho_{\{w,691\}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_{691}),$$

(i.e., related to Φ_w for $w = 2, 3, 4, \ldots$) all of them having the same semisimplication when reduced modulo 691. Specifically, its reduction $\bar{\rho}_{\{w,691\}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_{691})$ is a reducible representation and its semisimplication, $\bar{\rho}^{\mathrm{ss}}_{\{w,691\}}$, is equivalent to $\mathbf{1} \oplus \omega^{11}$. The appendix tells us that these $\rho_{\{w,691\}}$ all have the same preferred indecomposable residual representation, as well, and this representation has the property that the inertia group at 691 acts semisimply. William Stein and Craig Citro have calculated Φ_2, and its field of Fourier coefficient $\mathcal{F}_2$, and I am thankful to them for providing me with the information I'll be recounting here. Stein and Citro show that the (quotient) Hecke algebra—tensored with $\mathbf{Q}$— acting faithfully on the vector space of weight two cuspidal modular eigenforms of level 691 and nebentypus ω^{-10} is a field—and hence can be taken to be "our" $\mathcal{F}_2$. They show this field $\mathcal{F}_2$ to be of degree 57 over the cyclotomic field $\mathbf{Q}(\mu_{69})$; moreover, there is a (unique) prime ideal P of degree one, with residual characteristic 691 in the ring of integers of $\mathcal{F}_2$, such that if $\mathcal{F}_{2,v}$ is the completion of the field $\mathcal{F}_2$ with respect to the valuation determined by P, then under the canonical identification $\mathbf{Q}_{691} = \mathcal{F}_{2,v}$, for every $n \geq 1$, the n-th Hecke operator $T_n \in \mathcal{F}_2 \subset \mathcal{F}_{2,v} = \mathbf{Q}_{691}$ is equal to $a_2(n) \in \mathbf{Q}_{691}$.

We are this led to consider the abelian variety over $\mathbf{Q}$ that is simple (even over $\mathbf{C}$) related to the eigenform Φ_2. Call this abelian variety A for short (its standard name is $J_1(691; \omega^{-10})$). Explicitly, A is the abelian variety quotient of the jacobian of the modular curve $J_1(691)$ associated to $S_2(\Gamma_1(691; \omega^{-10})$, the space of weight two cuspidal modular eigenforms of level 691 and nebentypus

ω^{-10}. The calculations of Stein and Citro alluded to above give you that A is an abelian variety of a whopping dimension $2508 = 2 \cdot 22 \cdot 57$, whose endomorphism ring $\mathrm{End}_{\mathbf{C}}(A)$ tensored with $\mathbf{Q}$ is equal to the field $\mathcal{F}_2$. We know (see the appendix below) that its Galois module of 691-power torsion points has a subquotient of length two that is equivalent to $\bar{\rho}$.

In other words, $\bar{\rho}$ can be "found in the action of Galois on 691-torsion points of A." It is hard to believe that when we represent $\bar{\rho}$ in this manner we learn something[25] but we do![26]

Part III: The Wrench

13 A lemma in the style of Ribet

Here, in imprecise language, is the general question that the lemma addresses. Suppose that you have a continuous family of (Galois, say) group representations, the generic representation being irreducible. For argument's sake, suppose this family ρ_t is parametrized by a line, with parameter variable denoted t. For certain special values of t, say $t = t_o$ the representation ρ_{t_o} may no longer be irreducible. Borrowing the standard logo that algebraic geometers use to depict degeneration in a parametrized family of objects, we may sometimes think of our family of representations as depicted by the following cartoon[27].

[25] we actually learn at least two things: the first being, as mentioned, that we have an elegant occurrence of the representation $\bar{\rho}$ in the context of abelian varieties, but the second is that when we return to the 691-adic $G_{\mathbf{Q}}$-representation $\rho_{\Delta,691}$ on the two-dimensional vector space V over $\mathbf{Q}_{691}$, there is a unique choice of lattice, up to multiplication by a nontrivial scalar, for which $\bar{V}$ *is* a $G_{\mathbf{Q}}$-representation equivalent to $\bar{\rho}$ in $\boxed{\mathbf{4}}$. The argument for this latter statement is briefly given in the appendix.

[26] And to construct the abelian field $L^{\{200\}}$ (and the Galois representation $\rho^{\{200\}}$ that cuts it out) in a manner similar to the above, we would begin with the 691-adic cuspform $\Delta^{\{200\}}$ and would be discovering this Galois representation in Galois action on 691-power torsion in the abelian variety $J_1(691; \omega^{-198})$. William Stein informs me that this is a simple abelian variety of dimension 4928.

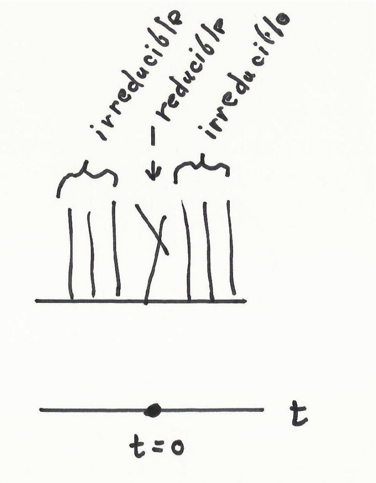
[27] A propos, note that *absolutely irreducibility* for degree two representations is an "open condition" in the following sense: Let A be an integral domain, G a group, and $\rho : G \to \mathrm{GL}_2(A)$ a homomorphism viewed as a family of G-representations $\rho_F : G \to \mathrm{GL}_2(F)$ varying over the collection of homomorphisms $A \to F$ for algebraically closed fields F. If ρ_{F_o} is reducible for any single injection $A \hookrightarrow F_o$, then ρ_F is reducible for all $A \to F$. The contrapositive, of course, implies that if ρ_F is absolutely irreducible for any single homomorphism $A \to F$ it is absolutely irreducible for F the field of fractions of A. **Proof:** Let M be a free A-module with G-action via ρ. Suppose that $M \otimes_A F_o$ is not irreducible, and that there admits a G-equivariant surjection $\phi : M \otimes_A F_o \to N$ to an F_o-vector space N of dimension one endowed with F_o-linear G-action. The action of G on N factors through an abelian quotient of G as does the action of G on $\ker\phi$. Now consider the restriction of ϕ to M and form

$$0 \to M_0 \to M \xrightarrow{\phi} M_1 \to 0$$

so that $M_1 \subset N$, and $M_0 \subset \ker\phi$. In particular, the actions of G on M_0 and M_1 factor through the abelian quotient of G. Now tensor with any field F to get the G-equivariant exact sequence

$$M_0 \otimes_A F \to M \otimes_A F \xrightarrow{\phi} M_1 \otimes_A F \to 0.$$

Since the actions of G on the two flanking modules factor through the abelian quotient of G it follows that $M \otimes_A F$ is reducible as degree two G-representation.



The Jordan-Hölder constituents of ρ_{t_o} are determined by the family $\{\rho_t\}_{t \neq t_o}$ parametrized by the complement of the point t_o. Nevertheless, as in algebraic geometric deformation theory—but not the less puzzling therefore— the isomorphism class of ρ_{t_o} is not completely determined: there may be many distinct ways of *filling in*[28] the family of representations $\{\rho_t\}_t$ at t_o. Often there is *one preferred* indecomposable representation. This is the lever that Ken uses so effectively.

Let G be a profinite group and $\mathcal{K}$ a complete discrete valued field with ring of integers $\mathcal{O}$, a choice of uniformizer π, and residue field denote $\kappa = \mathcal{O}/\pi\mathcal{O}$. To make a connection with the description of the previous paragraph you might, for example, take $\mathcal{K}$ to be the completion at t_o of the field of rational functions of one variable t over a topological field[29]κ. By definition an **$\mathcal{O}$-lattice** M in a finite dimensional $\mathcal{K}$-vector space V is a finitely generated $\mathcal{O}$-submodule of V such that the injection $M \hookrightarrow V$ induces an isomorphism $M \otimes_{\mathcal{O}} \mathcal{K} \cong V$. Note that the $\mathcal{O}$-lattice $\pi^{-m}\mathcal{O} \in \mathcal{K}$ is an open $\mathcal{O}$-submodule of $\mathcal{K}$, as is any $\mathcal{O}$-lattice in any finite dimensional $\mathcal{K}$-vector space (given the standard topology).

Ken's initial application was for $\mathcal{K} = \mathbf{Q}_p$ for some prime number p and $\kappa = \mathbf{F}_p$. We assume that κ is either a finite field or a field of characteristic 0, and begin with the following basic lemma.

Lemma 2. *Let G be a profinite group and V a finite dimensional $\mathcal{K}$-vector space endowed with a continuous $\mathcal{K}$-linear G-action, $r : G \to \mathrm{Aut}_{\mathcal{K}}(V)$.*

1. *There is an $\mathcal{O}$-lattice $M \subset V$ that is G-stable.*
2. *For each $g \in G$ the characteristic polynomial* $\det(1 - r(g)T)$ *of the action of g in the representation r has coefficients in $\mathcal{O}$.*

[28] An elegant way of treating this ambiguity in "filling in" is achieved by Wiles' notion of a *pseudo-character* or *pseudo-representation*; see for example the opening sections of [**?**] for an exposition, and also see section 1.1 of [**?**]. A very useful comprehensive discussion of the deformation theory of pseudo-representations (viewed as generalized determinants) and primer on the background of the subject—making extensive use of the earlier work of Procesi and others—can be found in [**?**].

[29] The field $\mathcal{K}$ inherits a topology from that of $\kappa = \mathcal{O}/\pi\mathcal{O}$ in the following way: for any $n \geq 0$, $\mathcal{O}/\pi^n\mathcal{O}$ is given the natural product topology inherited from that of κ in terms of the basis $\{1, \pi, \pi^2, \dots, \pi^{n-1}\}$; $\mathcal{O} = \lim_n \mathcal{O}/\pi^n\mathcal{O}$ is then given its projective limit topology and $\mathcal{K} = \cup_m \pi^{-m}\mathcal{O}$ its weak topology.

3. *If $M \subset V$ is a G-stable $\mathcal{O}$-lattice, form the associated "residual representation," $r_{M\otimes\kappa}$; i.e., the representation of G on the κ-vector space $M \otimes_{\mathcal{O}} \kappa$ obtained by the representation obtained from r via reduction $M \to M \otimes_{\mathcal{O}} \kappa$. Then for each $g \in G$ the characteristic polynomial*

$$\det(1 - r_{M\otimes\kappa}(g)T)$$

of the action of g in the representation $r_{M\otimes\kappa}$ is the reduction of $\det(1 - r(g)T)$ under the natural homomorphism $\mathcal{O}[T] \to \kappa[T]$.

4. *If $M, M' \subset V$ are two G-stable $\mathcal{O}$-lattices, the Jordan-Hölder constitutents[30] of the representations $r_{M\otimes\kappa}$ and $r_{M'\otimes\kappa}$ are equal.*

Proof: The proof of (1) starts with any lattice $M' \subset V$ and correcting it to be G-stable by noting that since the action $G \times M' \to V$ is continuous, M' is finitely generated over $\mathcal{O}$, and $M' \subset V$ is open, the subgroup $G_o \subset G$ stablizing M' is an open subgroup of G; moreover, since G is profinite, we have that $G_o \subset G$ is of finite index, and therefore G stabilizes $M := \sum g_i \cdot M' \subset V$ where $\{g_i\}_i$ is a (finite) system of representatives of left G_o-cosets in G. The proof of (2) and (3) are straightforward from this, while the final item (4), if κ is a finite field is an application of the classical Brauer-Nesbitt theorem (See [**?**] and also pp. 215-217, as well as Chapter XII, of [**?**]) while if κ is of characterstic 0 it is standard.

In view of this lemma, given a (finite dimensional) G-representation ρ on a $\mathcal{K}$-vector space we may speak of the various **residual representations attached to** ρ, these being representations of G into κ-vector spaces obtained as the reduction of the various G-stable $\mathcal{O}$-lattices in the underlying $\mathcal{K}$-vector space of the representation ρ. We may also speak of *the* **residual irreducible constituents of** ρ **with their multiplicities**, these being the irreducible representations that occur as Jordan-Hölder constituents of one (or, equivalently, any) *residual representation attached to* ρ.

Here is the lemma that is the backbone of Ken's strategy, expressed in general terms following Joel Bellaïche's article *A propos d'un lemme de Ribet* ([**?**]). We keep to the above terminology and hypotheses, fixing $\pi \in \mathcal{O}$ a uniformizer.

Lemma 3. (**Ribet-Bellaïche**) *Let G be a profinite group and V a finite dimensional $\mathcal{K}$-vector space endowed with a continuous irreducible $\mathcal{K}$-linear G-action, $r : G \to \mathrm{Aut}_{\mathcal{K}}(V)$. Let $\tilde{M} \subset V$ be a G-stable $\mathcal{O}$-lattice so that we may view r as a homomorphism $r_{\tilde{M}} : G \to \mathrm{Aut}_{\mathcal{O}}(\tilde{M})$ and let $r_{\tilde{M}\otimes\kappa} : G \to \mathrm{Aut}_{\kappa}(\tilde{M} \otimes_{\mathcal{O}} \kappa)$ be the corresponding residual representation.*

Now let $\bar{r}_0$ be a "proper" subquotient $\kappa[G]$-module in this residual representation $r_{\tilde{M}\otimes\kappa}$. "Proper" means that the degree of $\bar{r}_0$ (i.e., the dimension of the underlying κ-vector space) is positive and strictly less than the degree of r over $\mathcal{K}$ (equivalently, the degree of $r_{M\otimes\kappa}$ over κ).

Then there is a G-stable $\mathcal{O}$-lattice $M \subset V$ such that the associated residual representation $\bar{r} = r_{M\otimes\kappa}$ is isomorphic to a nonsplit extension of G-representations of κ-vector spaces, displayed here in terms of the labels of the corresponding representations:

[30] i.e., the irreducible representations occurring as subquotients of a Jordan-Hölder filtration, including their multiplicities,

$$0 \to \bar{r}_1 \longrightarrow \bar{r} \xrightarrow{\psi} \bar{r}_0 \to 0.$$

("Nonsplit" means that there is no $\kappa[G]$-equivariant homomorphism $\bar{r}_0 \to \bar{r}$ that is a left-inverse to ψ.)

Remark. Ribet's original result was formulated when $\bar{r}_1, \bar{r}_0$ above are characters, i.e., representations of degree one. A proof of this theorem for $\bar{r}_1, \bar{r}_0$ a pair of distinct absolutely irreducible representations of higher degree is due to Urban, in [**?**]. See also Theorem 1.1 of [**?**] for a different proof and see loc. cit. for remarks about the situation where $\bar{r}$ has more than two irreducible constituents.

Proof: By *lattice* (for short) we mean a G-stable $\mathcal{O}$-lattice in V. It is sometimes convenient to label the $\kappa[G]$-modules in the discussion below by the terms for the corresponding representations (e.g., $r_{\tilde{M}\otimes\kappa}$ and $\tilde{M}/\pi\tilde{M}$ are synonyms).

- **Step 1: The initial wrench, moving $\bar{r}_0$ to a quotient of the residual representation.**

  Since $\bar{r}_0$ occurs as a subquotient of the residual G-representation $\tilde{M} \otimes \kappa$ we may find a G-subrepresentation $\bar{N} \subset \tilde{M} \otimes \kappa$ such that $\bar{r}_0$ occurs as a quotient of $\bar{N}$. Let $M^{(0)} \subset \tilde{M}$ be the full inverse image in $\tilde{M}$ of $\bar{N} \subset \tilde{M} \otimes \kappa$ under the reduction homomorphism $\tilde{M} \to \tilde{M} \otimes \kappa$. Let $\bar{r}^{(0)}$ denote the residual representation $r_{M^{(0)}\otimes\kappa}$ associated to this new lattice $M^{(0)}$. We now have that $\bar{r}_0$ is a quotient of $\bar{r}^{(0)}$. Denote the kernel of the projection $M^{(0)} \otimes \kappa \to \bar{r}^{(0)}$ by $\bar{r}_1$. Consider the exact sequence of κ-representations

  $$0 \to \bar{r}_1 \to \bar{r}^{(0)} \to \bar{r}_0 \to 0. \qquad (\mathbf{1})$$

  Note that $\bar{r}_1$ is a representation of positive degree, given our hypothesis.

  If this exact sequence, $(\mathbf{1})$, of $\kappa[G]$-modules is nonsplit, then we are done, so—to continue—suppose that we have a splitting of $(\mathbf{1})$ (as G-representation). Fix such a splitting, i.e., a direct sum decomposition

  $$\bar{r}^{(0)} \;\simeq\; \bar{r}_1 \oplus \bar{r}_0.$$

- **Step 2: The inductive sequence of wrenches seeking a nonsplit extension.** Let $M^{(1)} \subset M^{(0)}$ be the full inverse image in $M^{(0)}$ of the subrepresentation $\bar{r}_0 \subset M^{(0)} \otimes \kappa$ under the reduction homomorphism $M^{(0)} \to M^{(0)} \otimes \kappa$. By construction, $M^{(1)}$ fits into two exact sequences of G-stable $\mathcal{O}$-modules,

  $$0 \to \pi M^{(0)} \to M^{(1)} \to \bar{r}_0 \to 0,$$

  and

  $$0 \to M^{(1)} \to M^{(0)} \to \bar{r}_1 \to 0. \qquad (\mathbf{2})$$

  Consider, now, the residual representation, $\bar{r}^{(1)} = r_{M^{(1)}\otimes\kappa}$, associated to this new lattice $M^{(1)}$.

  Lemma 4. *This $\kappa[G]$-representation $\bar{r}^{(1)}$ fits into an exact sequence*

  $$0 \to \bar{r}_1 \to \bar{r}^{(1)} \to \bar{r}_0 \to 0. \qquad (\mathbf{3})$$

Proof: By definition of $M^{(1)}$ we have a surjection $M^{(1)}/\pi M^{(1)} \to \bar{r}_0 \to 0$. Its kernel is canonically

$$\pi M^{(0)}/\pi M^{(1)} \cong M^{(0)}/M^{(1)} \cong \bar{r}_1,$$

the latter isomorphism by **(2)** above. □

Again, if this exact sequence (**3**) of $\kappa[G]$-modules is nonsplit, we are done. Otherwise we continue the same procedure as above, writing

$$\bar{r}^{(1)} \simeq \bar{r}_1 \oplus \bar{r}_0$$

(fixing a choice of splitting) and defining $M^{(2)} \subset M^{(1)}$ to be the full inverse image in $M^{(1)}$ of the subrepresentation $\bar{r}_0 \subset M^{(1)} \otimes \kappa$ under the reduction homomorphism $M^{(1)} \to M^{(1)} \otimes \kappa$. And again, $M^{(2)}$ fits into two exact sequences of G-stable $\mathcal{O}$-modules,

$$0 \to \pi M^{(1)} \to M^{(2)} \to \bar{r}_0 \to 0,$$

and

$$0 \to M^{(2)} \to M^{(1)} \to \bar{r}_1 \to 0.$$

This gives us four things:

1. an inclusion $\pi^2 M^{(0)} \subset M^{(2)}$ with quotient an extension of $\bar{r}_0 = M^{(2)}/\pi M^{(1)}$ by

   $$\pi M^{(1)}/\pi^2 M^{(0)} \cong M^{(1)}/\pi M^{(0)} \cong \bar{r}_0,$$

   i.e., the quotient $\mathcal{R}_0^{(2)} := M^{(2)}/\pi^2 M^{(0)}$ is an extension of $\bar{r}_0$ by $\bar{r}_0$, and

2. an inclusion

   $$M^{(2)} \subset M^{(0)}$$

   with quotient $\mathcal{R}_1^{(2)} := M^{(0)}/M^{(2)}$ being an extension of $\bar{r}_1$ by $\bar{r}_1$.

3. $$M^{(2)} + \pi M^{(0)} = M^{(1)} \subset M^{(0)}.$$

4. Combining these we would get an exact sequence

   $$0 \to \mathcal{R}_0^{(2)} \to M^{(0)}/\pi^2 M^{(0)} \to \mathcal{R}_1^{(2)} \to 0. \quad (\mathbf{4})$$

   Note that $\mathcal{R}_0^{(2)} \otimes_{\mathcal{O}} \kappa \cong \bar{r}_0$ and that $\mathcal{R}_1^{(2)} \otimes_{\mathcal{O}} \kappa \cong \bar{r}_1$, and tensoring (**4**) over $\mathcal{O}$ with κ yields an exact sequence

   $$0 \to \bar{r}_0 \to \bar{r}^{(1)} \to \bar{r}_1 \to 0.$$

- **Step 3: An infinite sequence of sublattices**

  We discover in this manner that either we are done at some stage, or else we have an infinite sequence of G-stable sublattices

  $$\cdots \subset M^{(i+1)} \subset M^{(i+1)} \subset \cdots \subset M^{(0)}$$

  with the following properties.

1. We have an exact sequence
$$0 \to \pi^i M^{(0)} \to M^{(i)} \to \mathcal{R}_0^{(i)} \to 0,$$
where $\mathcal{R}_0^{(i)}$ has a Jordan-Hölder filtration of length i with successive quotients isomorphic to $\bar{r}_0$.
2. For $j \le i$ we have that
$$M^{(i)} + \pi^j M^{(0)} \ = M^{(j)} \subset M^{(0)}$$
3. We have an exact sequence
$$0 \to M^{(i)} \to M^{(0)} \to \mathcal{R}_1^{(i)} \to 0,$$
where $\mathcal{R}_1^{(i)}$ has a filtration of length i with successive quotients isomorphic to $\bar{r}_1$.
4. We have an exact sequence
$$0 \to \mathcal{R}_0^{(i)} \to M^{(0)}/\pi^i M^{(0)} \to \mathcal{R}_1^{(i)} \to 0, \quad (\mathbf{5i})$$
and tensoring ($\mathbf{5}$) over $\mathcal{O}$ with $\mathcal{O}/\pi^j\mathcal{O}$ for any $j \le i$ yields the exact sequence
$$0 \to \mathcal{R}_0^{(j)} \to M^{(0)}/\pi^j M^{(0)} \to \mathcal{R}_1^{(j)} \to 0. \quad (\mathbf{5j})$$

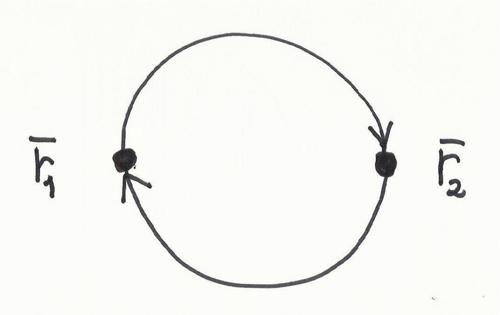
From (1) we get the inclusion
$$\mathcal{R}_0^{(i)} \subset M^{(0)}/\pi^i M^{(0)}$$
while from (2) we get that the natural projections $M^{(0)}/\pi^{i+1}M^{(0)} \to M^{(0)}/\pi^i M^{(0)}$ induce surjections
$$\mathcal{R}_0^{(i+1)} \to \mathcal{R}_0^{(i)}.$$
From (3) we get surjections
$$\mathcal{R}_1^{(i+1)} \to \mathcal{R}_1^{(i)}.$$
Passing to the projective limits,
$$\mathcal{R}_0 := \lim_i \mathcal{R}_0^{(i)} \subset \lim_i M^{(0)}/\pi^{i+1}M^{(0)} \ = \ M^{(0)}$$
and
$$\mathcal{R}_1 := \lim_i \mathcal{R}_1^{(i)}$$
we get an exact sequence of G-stable $\mathcal{O}$-modules $0 \to \mathcal{R}_0 \to M^{(0)} \to \mathcal{R}_1 \to 0$. Under our assumption that $\bar{r}_1$ is proper, we get that both $\mathcal{R}_0$ and $\mathcal{R}_1$ are of infinite length—i.e., of positive rank—as $\mathcal{O}$-modules. Tensoring with $\mathcal{K}$ we then get that our original K-representation r has a proper G-stable subspace, contradicting the fact that it is assumed to be irreducible.

□

A directed graph: Let $r : G \to \mathrm{Aut}_{\mathcal{K}}(V)$ be an irreducible representation as in the lemma, and let $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_\nu$ be the irreducible residual constituents of r. Suppose for simplicity of notation that the $\bar{r}_i$ are all distinct—i.e., they each occur with multiplicity one in $r_{M\otimes\kappa}$ for any G-stable $\mathcal{O}$-lattice $M \subset V$. Construct a graph $Y := Y(r)$ as follows. The vertices of Y are the residual constituents $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_\nu$. Draw a "directed edge" from $\bar{r}_i$ to $\bar{r}_j$ if there is a G-stable $\mathcal{O}$-lattice $M \subset V$ such that the associated residual representation $r_{M\otimes\kappa}$ possesses a subquotient which is a *nontrivial* extension of $\bar{r}_i$ by $\bar{r}_j$. Using the statement and proof of the above lemma a further argument will show that the graph Y is connected and each vertex of Y has as least one directed edge leaving it, and another entering it[31]. (It is easy enough to make examples where the graph is given by a single (directed) cycle lacing through all the residual irreducible constituents.)

The immediate effect of Ribet's strategy is to provide a large quantity of nontrivial extensions. The surprise is that (in many cases) one has—by application of this strategy—constructed extensions with very precise and useful further properties.

A major (but not the only) application of this lemma is when the residual representation has two distinct irreducible contitutents, $\bar{r}_1$ by $\bar{r}_2$, each occurring with multiplicity one. In this situation one obtains, given the hypotheses of Lemma **??**, two indecomposable residual representations: a nontrivial extension of $\bar{r}_1$ by $\bar{r}_2$ and a nontrivial extension of $\bar{r}_2$ by $\bar{r}_1$, giving us this elementary "complete" graph on two vertices[32].



A well known classical example of this is when $p = 5$, $\mathcal{K} = \mathbf{Q}_5, \kappa = \mathbf{F}_5$, and $G = G_{\mathbf{Q},\{5,11\}}$ with the representation $r : G_{\mathbf{Q},\{5,11\}} \to \mathrm{GL}_2(\mathbf{Q}_5)$ coming from the action of Galois on the 5-power torsion subgroup of the elliptic curve $X_0(11)$. The two indecomposable residual representations are the representations of Galois on the 5-torsion of the two other elliptic curves over $\mathbf{Q}$ of conductor 11.

[31] If Y is expressible as a disjoint union of two nonempty subgraphs, one first shows that for any lattice M the associated residual representation $\bar{r}$ has a direct sum decomposition $\bar{r} = \bar{r}_1 \oplus \bar{r}_0$ where each of the summands $\bar{r}_1, \bar{r}_0$ has irreducible constituents corresponding to the vertices of each of the two subgraphs. Moreover, under this hypothesis, for *every* lattice for which the associated residual representation $\bar{r}$ has a quotient representation isomorphic to $\bar{r}_0$ we have such a direct sum representation. The technique of the proof of Lemma **??** then works to show that this is impossible, thanks to the irreducibility of the $\mathcal{K}$-representation r. To show that each vertex of the graph has an edge directed to it (as well as emerging from it) apply Lemma **??** appropriately to the dual of r.

[32] Here is the argument that these nontrivial extensions, as elements of $Ext_{k[G]}(\bar{r}_1, \bar{r}_2)$ and $Ext_{k[G]}(\bar{r}_2, \bar{r}_1)$, are independent of the lattice for which they are residual representations: suppose that you have two lattices M, M' with indecomposable residual representations, $\bar{r}, \bar{r}'$, both admitting, say, a $k[G]$-equivariant surjection to $\bar{r}_2$. Multiplying by an appropriate power of π you can arrange it so that $M \subset M'$ but M is not contained in $\pi M'$. Then *either* $M = M'$ *or else* the image of M in $\bar{r}'$ is isomorphic to $\bar{r}_2$, thereby splitting $\bar{r}'$ contrary to hypothesis.

14 Liminal representations

One can also consider the prospect of “going the other way.” That is, given a reducible representation ρ_0 of G in a finite dimensional κ-vector space, when can we find a (locally analytic, say) family of representations ρ_t parametrized by a variable t ranging through a disc about the origin such that

- the generic member of the family is an *irreducible* representation of G, and
- the specialization of this family to $t = 0$ is our initial representation ρ_0?

If the above happens, let us say that the representation ρ_0 is **limit-irreducible**, or for short, call it **liminal**.

When we say that a given representation $\rho : \mathcal{G}_{K,S} \to \mathrm{GL}_2(\mathbf{Z}_p)$ *is the limit of representations satisfying a particular property* P we mean, more explicitly, that there is a sequence $\rho_i : \mathcal{G}_{K,S} \to \mathrm{GL}_2(\mathbf{Z}_p)$ of continuous homomorphisms each satisfying property P (for $i = 1, 2, 3, \dots$) converging in the p-adic topology to ρ. So *liminality*—in this set-up—amounts to being a member of a locally analytic family of representations that is the limit of a sequence *irreducible* representations in the family. We will usually also want our family ρ_t to have some specifically described further properties. We will presently see that, in certain good situations, we can predict that the liminal representation, when restricted to the decomposition group at p, has surprisingly good properties[33].

A (technically only slightly) different version of this concept of *liminality* is when there a complete discrete valued field $\mathcal{K}$, as in the previous discussion, with residue field κ, and for which there is a continuous *irreducible* representation ρ of G into a finite-dimensional $\mathcal{K}$-vector space having ρ_0 as one of its residual κ-vector space representations.

Part IV: Types of p-adic Galois Representations of degree two

15 Ordinary and Nearly Ordinary Galois representations

Let K be a number field, $\bar{K}/K$ an algebraic closure and put $\mathcal{G}_K := \mathrm{Gal}(\bar{K}/K)$. Let S be a finite set of places of K let $\mathcal{G}_{K,S}$ be the quotient of $\mathcal{G}_K$ obtained by dividing by the closed normal subgroup generated by all inertia groups for places outside K. Fix p a prime, let $S(K,p)$ denote the set of all places of K dividing p, and suppose that S contains $S(K,p)$.

[33] In a more general–yet still particular–context, a big role in recent developments is played by a theorem of Kisin that allows us to deduce that liminal representation ρ possesses a limiting *p-adic period* in the sense of Fontaine’s theory, provided the approximating representations ρ_i, when restricted to the decomposition group at p, satisfy appropriate requirements (see [**?**] and section 3 of [**?**]).

We will now formulate a sequence of conditions on degree two Galois Representations. Let M be a free $\mathbf{Z}_p$-module of rank two endowed with a continuous $\mathbf{Z}_p$-linear action of $\mathcal{G}_{K,S}$. Let

$$\rho : \mathcal{G}_{K,S} \longrightarrow \mathrm{Aut}(M) \simeq \mathrm{GL}_2(\mathbf{Z}_p)$$

denote the corresponding degree two representation.

Definition 1.

1. *We say that ρ is* **nearly ordinary** *if for all $v \in S(K,p)$ the restriction ρ_v of ρ to a decomposition group $\mathcal{G}_{K_v}$ at v preserves a free rank one sub-module $M_v \subset M$. We can, and do, take M_v to be saturated in M; we choose a decomposition subgroup for v and such an M_v for each $v \in S(K,p)$ and—in cases where there are more than one such possible M_v—we view these choices as part of the nearly ordinary structure of ρ; cf. [C-M]. We then have for each $v \in S(K,p)$ an exact sequence of $\mathcal{G}_{K_v}$-modules*

$$0 \to M_v \to M \to M/M_v \to 0.$$

   The restriction of ρ_v to M_v and the induced action on M/M_v give us two degree one characters, i.e., homomorphisms $\mathcal{G}_{K_v} \to \mathbf{Z}_p^*$, for each $v \in S(K,p)$. For evident reasons let us call the character giving the action on M_v the **local sub-character of the nearly ordinary representation ρ at** *v; and call the character giving the action on M/M_v the* **local quotient-character***.*

2. *We say that a nearly ordinary (degree two) Galois representation is* **ordinary** *if the local quotient-characters of ρ (for $v \in S(K,p)$) are all unramified*[34]*.*

3. *We say that an ordinary (degree two) Galois representation is* **anomalous** *if the local quotient characters of ρ (for $v \in S(K,p)$) are all trivial.*

For reasons that will become apparent below, we will define *Iwasawa representations* to allow for more general scalars than the $\mathbf{Q}_p$ that was operative in the above definition. For brevity, then, let us consider a continuous Galois (G_K) representation on a two-dimensional vector space V over $\mathbf{C}_p$; $\rho : G_K \to \mathrm{Aut}_{\mathbf{C}_p}(V) \approx \mathrm{GL}_2(\mathbf{C}_p)$.

Definition 2. *We will call such a representation an* **Iwasawa representation** *if*

1. *the global representation ρ is indecomposable,*

2. *the semisimplication of ρ is the sum of two characters of G_K, a nontrivial character $\chi : G_K \to \mathbf{C}_p^*$ and the trivial character $\mathbf{1}$,*

3. *the character $\chi = \det(\rho)$ has the property that its minimal splitting field over K is a $\mathbf{Z}_p$-extension of K.*

[34] There is a somewhat arbitrary choice to be made here: some texts define *ordinary* by the requirement that the local *sub*-characters of ρ (for $v \in S(K,p)$) are all unramified. The two choices are elementarily related in that if a $\mathbf{Z}_p[\mathcal{G}_K]$-module M is ordinary according to one of these choices, its $\mathbf{Z}_p$-dual, $Hom(M, \mathbf{Z}_p)$, will be ordinary according to the other. In effect, if you are dealing with cohomology, the choice we've just made is slightly more natural than the other, and if you are dealing with homology the situation is reversed.

4. *the character χ occurs as a sub-representation of ρ and the trivial character* $\mathbf{1}$ *occurs as a quotient-representation,*

5. *for all $v \in S(K,p)$ the local representation ρ_v when restricted to the inertia group $\mathcal{I}_{K_v} \subset \mathcal{G}_{K_v}$ is semisimple.*

Note that if an Iwasawa representation takes its values in $\mathrm{GL}_2(\mathbf{Z}_p)$ we may view it as a nearly ordinary representation, and indeed—for each $v \in S(K,p)$— we have our choice as to which of the two local degree one characters we choose to be the sub-character, which the quotient-character. Choosing, consistently, the local quotient-character to be the trivial character as we can do by (5), the Iwasawa representation is then anomalous.

16 The wrench: guaranteeing that certain abelian extensions are unramified at p

At this point we will finally pick up the issue hinted at in the $\boxed{\mathbf{4}} \Rightarrow \boxed{\mathbf{3}}$ lap in the discussion-of-proof at the end of Part I above. Why is the extension described in $\boxed{\mathbf{3}}$ everywhere unramified?

Our representation $\bar{\rho}$ is unramified at primes ℓ different from p since $\bar{\rho}$ is contained in the Galois representation on p-torsion in the abelian variety $J_1(p)/J_0(p)$ which has good reduction outside the prime p (see the discussion in section **??**)[35]. To show that $\bar{\rho}$ is *everywhere unramified* we must show it to be unramified at the delicate prime p.

This follows from the *wrench phenomenon* and I will illustrate it in two distinct situations.

- **(Ribet's Context)** The following wrench is merely a toy illustration, but the basic idea has an analogue in the full generality required by Ribet.

  Suppose that you have an elliptic curve E over $\mathbf{Q}$ with good reduction at an odd prime number p, and the elliptic curve has a rational point of order p. (This—to be sure—doesn't happen for too many primes p, but **(a)** it *does* happen and **(b)** —as I said—this is a toy illustration.) A perfectly good example, already mentioned at the end of section **??**, is $E = X_1(11)$ and $p = 5$.

  Now when you have such an elliptic curve, consider its $G_{\mathbf{Q}} := \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$-module of p-torsion points $E[p]$, which is a vector space of dimension two over $\mathbf{F}_p$. This vector space will be wrenched in two contrary ways.

  - First, as a global Galois representation, it fits into an exact sequence

    $$0 \to \mathbf{Z}/p\mathbf{Z} \to E[p] \to \mu_p \to 0$$

    where μ_p is the group of p-th roots of unity, viewed as degree one Galois representation over $\mathbf{F}_p$.

[35] This would also follow from our discussion regarding $\bar{\rho}$ as related to Deligne's theorem (See section **??**) since, for all $\ell \neq p$ Deligne obtains the representation from the Galois representation on the mod p étale cohomology of a smooth abelian variety over $\mathbf{F}_\ell$.

- But (for finite group scheme reasons) if you restrict this representation to $D_p \subset G_{\mathbf{Q}}$, a decomposition group at p, you find it fitting into an exact sequence where the $\mathbf{Z}/p\mathbf{Z}$ and μ_p occur in the opposite order:

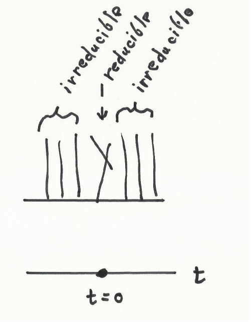
$$0 \to \mu_p \to E[p] \to \mathbf{Z}/p\mathbf{Z} \to 0.$$

Since the D_p-representations $\mathbf{Z}/p\mathbf{Z}$ and μ_p are distinct, we learn from this *wrench* that D_p acts semisimply on $E[p]$ (this D_p-representation being then just the direct sum $\mathbf{Z}/p\mathbf{Z} \oplus \mu_p$ and therefore is split by the cyclotomic field $K = \mathbf{Q}(\mu_p) = \mathbf{Q}(e^{2\pi i/p})$. In particular, if $L/\mathbf{Q}$ is the splitting field (over $\mathbf{Q}$) of the $G_{\mathbf{Q}}$-representation $E[p]$ then L/K is *unramified at* p. If $E[p]$ is also indecomposable as a $G_{\mathbf{Q}}$-representation, the L/K will be a cyclic degree p representation unramified at p. In our toy example, $E = X_1(11)$ and $p = 5$ we have hereby constructed a cyclic degree 5 extension of $\mathbf{Q}(e^{2\pi i/5})$ that is unramified at the unique prime above 5 and only ramified at primes above 11.

Ribet employs this argument *not* for elliptic curves as such but rather for abelian varieties over $\mathbf{Q}$ of the form $J_1(p)/J_0(p)$ (or abelian varieties isogenous to quotients of $J_1(p)/J_0(p)$) these having good reduction outside of p. Therefore, ramification at p is the only issue in question here. But these abelian schemes actually achieve good reduction also at p over the field $\mathbf{Q}(\mu_p)^+$ (cf. [**?**]) which is enough to press an analogue of the above argument forward[36].

- **(Parameter spaces of ordinary degree two Galois representations)**

Let us return to the context of the beginning of section **??** and suppose you are given a family ρ_t of degree two *ordinary* p-adic $G_{\mathbf{Q}}$-representations parametrized by an open disc in $\mathbf{Q}_p$, with parameter variable denoted t.



Suppose that these ρ_t's are generically irreducible in the sense described previously, and for the value $t = t_o$ the representation ρ_{t_o} is not irreducible—and therefore even though its *semisimplification* is well defined—as previously discussed—it itself is not well-defined; suppose that it has as its Jordan-Hölder constituents the $G_{\mathbf{Q}}$-characters $\mathbf{1}$ and χ where χ is ramified at p. We may "fill in" the family in more than one way; we choose ρ_{t_o} to be indecomposable and such that we have an exact sequence of $G_{\mathbf{Q}}$-representations with the following the "ordering of characters":

$$0 \to \mathbf{1} \to \rho_{t_o} \to \chi \to 0.$$

[36] More specifically, working with such an abelian scheme locally over the completion of $\mathbf{Z}(\mu_p)^+$ at the prime above p one uses a crucial result of M. Raynaud ([**?**]) that gives information about the structure of finite flat sub-group schemes of exponent p over discrete valuation rings that are finite degree extensions of $\mathbf{Z}_p$ of ramification index $< p-1$ together with a wrenching argument very similar to the one in the previous paragraph to conclude.

Now, since our family is ordinary, and χ is ramified at p, the D_p-representation obtained from ρ_{t_o} also fits into an exact sequence of D_p-representations, but with the opposite ordering:

$$0 \to \chi \to \rho_{t_o} \to \mathbf{1} \to 0.$$

The same *wrench phenomenon applies* as in the previous bullet to give us that if K is the splitting field of χ over $\mathbf{Q}$, and L is the splitting field of ρ_{t_o} over $\mathbf{Q}$, then $K \subset L$ and L/K is unramified at p (interpreted appropriately, since these in general will be extension fields of infinite degree over $\mathbf{Q}$).

17 Liminality of Iwasawa representations

The proof of the main conjecture ([**?**] over $\mathbf{Q}$; and more generally, [**?**] over totally real fields) tells us something (not yet everything!) about liminality of Iwasawa representations. Specifically, in the special case over $\mathbf{Q}$, for χ a p-adic character of $G_{\mathbf{Q}}$ let $\chi^* = \epsilon\chi^{-1}$ where $\epsilon : G_{\mathbf{Q}} \to \mathbf{Z}_p^*$ is the p-cyclotomic character.

1. The p-adic Leopoldt-Kubota L-function, $L_p(\xi)$, vanishes at the p-adic character $\xi = \chi^*$ if and only if there is an Iwasawa representation of determinant χ.
2. There is one and only one Iwasawa representation of determinant χ if and only if the zero of L_p is simple at χ^* and in this case the Iwasawa representation occurs an indecomposable representation ρ_s attached to a point s on a one-parameter p-adic analytic family (technically: a *Hida family*) of (generically irreducible) degree two Galois representations.

A consequence is that if the zero of L_p is simple at χ^* then the corresponding Iwasawa representation is indeed liminal in a very strong way: it is a limit of geometric, ordinary, irreducible representations[37].

Sidenote: There is something baffling happening in the (unlikely) event that the Leopoldt-Kubota L-function has a multiple zero at χ^*. For then we have a positive dimensional projective space parametrizing all Iwasawa representations of determinant χ, and yet only finitely many of these will lie in Hida components. Which ones? Or are all zeroes the Leopoldt-Kubota L-function simple?

Part V: Liminality in recent and current work

[37] If K is a totally real field, (and—for simplicity of discussion—assume that Leopoldt's conjecture holds for K) by the proof of the main conjecture for K (cf. [W]) one sees that the same statement holds for Iwasawa representations over K.

18 The general framework

To succinctly remind ourselves of Ribet's idea—but framing it in the more general context of reductive groups–we may illustrate the procedure by these five steps (allowing for variants, the most evident variant being a *congruence version* of what we describe below, such as in the format originally used by Ribet himself [38]).

1. *The equipment:*

   Let p be a prime number, and K a number field.

   Let $H^{(1)}, H^{(2)}, \dots, H^{(\nu)}$ and G be a collection of reductive groups over $\bar{\mathbf{Q}}_p$, and let

   $$\rho^{(i)} : G_K \to H^{(i)}(\bar{\mathbf{Q}}_p)$$

   $(i = 1, 2, \dots, \nu)$ be irreducible Galois representations[39]. In practice, these groups will be either general linear, symplectic, or unitary groups and we will be considering them as subgroups of general linear groups via their standard representations. These Galois representations will, in recent practice again, either be of degree one—i.e., characters—or, more generally, will be obtained from automorphic forms for reductive groups of symplectic or unitary type over $\mathbf{Q}$, or over totally real or CM number fields. We will call these $\rho^{(i)}$ the *constituent representations.*

   Let $P \subset G$ be a parabolic subgroup with $P = H \cdot U$ a Levi decomposition, where $H = H^{(1)} \times H^{(2)} \times \cdots \times H^{(\nu)}$. The goal is to find interesting indecomposable G_K representations that are extensions of pairs of constituent representations. This will be done by "constructing" G_K-representations ρ_o into $P(\bar{\mathbf{Q}}_p) \subset G(\bar{\mathbf{Q}}_p)$ with the $\rho^{(i)}$ as Jordan-Hölder constituents, and then pick out appropriate two-stage subquotient representations. To this aim, one considers some *well chosen* reductive group $\mathcal{G}$ over a number field admitting automorphic forms that have associated Galois representations $G_K \to G(\bar{\mathbf{Q}}_p)$. We call the reductive group $\mathcal{G}$ "the" *encompassing reductive group.*

2. *The semi-simplification of the indecomposable representation to be constructed is related to automorphic forms:* Define

   $$\rho_o^{ss} : G_K \to H(\bar{\mathbf{Q}}_p) = \prod_{i=1}^{\nu} H^{(i)}(\bar{\mathbf{Q}}_p)$$

   to be the product representation. The first real step in this process is to find an automorphic representation π_o for the encompassing reductive group $\mathcal{G}$ that has an associated Galois representation whose semi-simplification is ρ_o^{ss}. In the initial use, and in early applications, this automorphic form on the encompassing reductive group was taken to be an appropriate Eisenstein series, but certain other nontempered representations have also been brought into play, and the all-important passage from $\boxed{\mathbf{6}}$ to $\boxed{\mathbf{5}}$ in Ribet's original method, which bridges–via the convenience of congruences—the divide between *Eisenstein series* and *cuspforms*, is sometimes replaced by bridging the divide between *nontempered* and *tempered* automorphic representations, following a suggestion made many years ago by Harder.

[38] also when unitary groups are involved, and you have a quadratic extension K/k in the works to contend with

[39] where by *irreducible* all I mean is that the composition of these representations with the standard representation yields an irreducible representation of G_K into the corresponding general linear group.

3. *The automorphic form is "fit" into a p-adic family:* "Fit" π_o into a p-adically interpolable family of automorphic representations of the encompassing reductive group, and prove that there is indeed an associated (rigid analytic) family of Galois representations

$$\rho_t : G_K \to G(\bar{\mathbf{Q}}_p)$$

parametrized by an irreducible ("pointed") rigid analytic space (T, t_o) such that

- for $t = t_o$, the semisimplification of ρ_{t_o} is ρ_o^{ss},
- for $t \neq t_o$ the ρ_t are irreducible representations,
- for a Zariski-dense set of values of t, the ρ_t's have good properties.

4. *The wrench is used to get a desired liminal representation:* Now use the lemma of Ribet-Bellaïche and the "wrench phenomenon" described earlier to modify, if necessary, the limiting ρ_{t_o} so as to get an interesting liminal Galois representation—call it ρ_o. The number field cut out by the *nonsemisimple* representation ρ_o constructs, for us, Galois extensions of the number field cut out by the original constituent representations—described in the Ribet-Bellaïche lemma.

5. *Establishing good properties of the action of decomposition group:* Finally, use whatever good properties the ρ_t have when restricted to the decomposition groups at places above p to guarantee good local behavior (at places above p) for ρ_o and hence for the nontrivial extensions of the constituent representations that you have constructed, thereby showing that certain of these extensions provide sought-for elements in appropriate Selmer groups.

Here are a few examples to give a brief rough idea of the type of work that has been done[40], and that is being done, along these lines[41]. Below, the symbol χ will just mean some Galois character to $\bar{\mathbf{Q}}_p^* = \mathrm{GL}_1(\bar{\mathbf{Q}}_p)$ but, in practice, one may also want to deal with characters to multiplicative groups of extension fields of $\bar{\mathbf{Q}}_p$.

1. The above seems not to be too bad a strategy if you want to prove "main conjectures" as in [**?**] (the main conjecture of Iwasawa theory over $\mathbf{Q}$, and more generally as in [**?**] (the main conjecture of Iwasawa theory over totally real fields). The large difference between the approaches in [**?**] and [**?**] is that (although both follow the Ribet wrench philosophy) [**?**] makes extensive and particular use of the algebraic geometric structure of the jacobian of modular curves (which is not available in a more general setting) while [**?**] replaces this with the more automorphic format as described above. Here $\nu = 2$, the reductive groups $H^{(1)}, H^{(2)}$ are both isomorphic to GL_1, the constituent representations are two characters, one of them the trivial character. The encompassing reductive group is GL_2 and the parabolic subgroup $P \subset G$ is the Borel subgroup of upper triangular matrices. We can summarize this by the picture of the 2×2-matrix that we'll think of as a $(1+1) \times (1+1)$ matrix:

$$\begin{pmatrix} 1 & * \\ 0 & \chi \end{pmatrix}$$

[40] It is noteworthy—and natural— that many, perhaps all, of the classic results proving modularity of two-dimensional representation of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ (that satisfy appropriate hypotheses) make use of procedures that touch on Ribet's wrench; to cite one beautiful example, see [**?**].

[41] I am thankful to Michael Harris for help in preparing this brief summary.

the extension constructed (signified by the "$*$ in the upper right-hand corner") providing us with elements in the one-dimensional Galois cohomology of the degree one representation χ and, thanks to Step 5 above, these elements enjoy good local properties. The *congruence version* of this is, of course, the strategy initiated by Ribet. An appropriate Hida family gives the p-adically varying family of Galois representations "ρ_t."

2. To construct elements in the Selmer group of an adjoint representation of a Galois representation $\rho : G_K \to \mathrm{GL}_d(\bar{\mathbf{Q}}_p)$ one might try the above strategy with ρ and $\rho^* \otimes \chi$ as constituent characters, with $G = \mathrm{GSP}_{2d}$ where the picture corresponding to the one drawn in (2), just above, is of the $2d \times 2d$ matrix:

$$\begin{pmatrix} \rho & * \\ 0 & \rho^* \otimes \chi \end{pmatrix}$$

and the parabolic subgroup $P \subset G$ is the evident one. This is the format described in 1997 by Haruzo Hida, Jacques Tilouine, and Eric Urban [**?**] in their strategy for a possible proof of the main conjecture for the adjoint representation $ad(\phi)$ where ϕ is the GL_2 automorphic representation attached to an elliptic curve over $\mathbf{Q}$. Here $d = 2$, $G = \mathrm{GSP}_4$, and one then may hope to use p-adic rigid analytic families of GSP_4-automorphic forms to effect Steps (2) and (3) of the strategy outlined above, and thereby to produce elements in the Selmer group of twists of the adjoint representation to ρ. This was formulated as a candidate strategy in [**?**] and many of the technical hurdles to carry it out were dealt with in that article. As Michael Harris has explained to me, what was principally left yet to be done (given [**?**]) to obtain this main conjecture was to guarantee non-divisibility by p of (non-constant) Fourier coefficients of certain Eisenstein series.

One has, in this story of the main conjecture for the adjoint representation $ad(\phi)$, three basic objects: the p-adic L-function interpolating the critical values of the symmetric square of the modular forms in these families, the characteristic ideal of the associated Selmer group, and a characteristic *Eisenstein* ideal containing information on the congruences between cuspidal Siegel modular forms of genus 2 and the Klingen-type Eisenstein series. Regarding this, see [**?**] where, in an appropriate context, the divisibility of the Eisenstein ideal alluded to above by the L-function is shown, and see [**?**] for the divisibility of the characteristic ideal of the Selmer group by the Eisenstein ideal.

3. Finding elements of the Selmer group when the sign of the functional equation would predict that they exist seems to be amenable to the above outlined approach.

   (a) A "congruence version" of **(1)**—i.e., a strategy close to Ribet's original strategy—was carried out by Joel Bellaïche for quadratic imaginary fields in his thesis [**?**] (cf., the recently published [**?**]) where for a set of primes p of positive density, a Ribet-type theorem was proven relating nontriviality of the p-primary group of the Selmer group of algebraic Hecke characters over an imaginary quadratic field if the sign of the corresponding L-function is -1. As alluded to above, Harder's suggestion (of replacing the Eisenstein series that played the role —in $\boxed{\mathbf{6}}$—of Ribet's proof by a CAP automorphic form is employed by Bellaïche in his approach, coupled with an idea due to Clozel, Bellaïche's thesis advisor[42]. Here $\nu = 3$, the encompassing reductive group is a unitary group in three variables, and if η is the algebraic Hecke character one is studying, the residual representation attached to the"ρ_o^{ss}" of the above method is a sum of three characters with the residual representation attached to η occurring as one of the constituents.

[42] See the discussion in section 1 of [**?**] related to [**?**].

(b) Starting with a self-dual $G_{\mathbf{Q}}$-representation representation ρ into $\mathrm{GL}_2(\bar{\mathbf{Q}}_p)$, and a character χ, and again letting $\nu = 3$, consider the triple of constituent $G_{\mathbf{Q}}$-representations χ, ρ, χ^* (with values in $\mathrm{GL}_1(\bar{\mathbf{Q}}_p), GL_2(\bar{\mathbf{Q}}_p), GL_1(\bar{\mathbf{Q}}_p)$ respectively) and take $G = \mathrm{GSP}_4$. Here—if the sign is right— the aim would be to find $G_{\mathbf{Q}}$-representations of the shape given by the following picture of this $(1+2+1) \times (1+2+1)$ matrix:

$$\begin{pmatrix} \chi & 0 & * \\ * & \rho & * \\ * & 0 & \chi^* \end{pmatrix}$$

whose most "usable" pieces are the submatrices:

$$\begin{pmatrix} \chi & 0 \\ * & \rho \end{pmatrix}$$

and

$$\begin{pmatrix} \rho & * \\ 0 & \chi^* \end{pmatrix}$$

which provide the sought-for extension(s) of Galois representations (these two being equivalent under duality).

This is the format of the article [**?**] where ρ is taken to be a $G_{\mathbf{Q}}$-representation attached to a newform of (even) weight ≥ 2 for $\Gamma_0(N)$ (some N) where the functional equation for the L-function would predict that the Selmer group associated to ρ is of *odd* rank. In this case Skinner and Urban prove that the rank is, at least, positive. As in Bellaïche's thesis, the Eisenstein series (of $\boxed{\mathbf{6}}$ of Ribet's proof) is replaced by a CAP form. Here, as above, *non-ordinary* p-adic deformations are used to obtain the desired element in the Selmer group.

4. A variant of **(4)** above is to keep to the same $(1+d+1) \times (1+d+1)$ matrix picture (for various values of d) but using a unitary group of rank $d+2$ rather than a (general) symplectic group. Here one would work with initial representations χ, ρ, χ^* in that order, with χ some appropriate character and ρ a Galois representation associated with an automorphic form for a unitary group of rank d. Chenevier's thesis adopts such a format, and—taking $d = 1$, E a quadratic imaginary field and a unitary group of rank three, the joint work of Bellaïche and Chenevier [**?**] employs this to find elements in the Selmer elements of certain algebraic Hecke characters over E when the functional equation sign would predict that such elements should exist. Here a nontempered π_o is used. The forthcoming volume of Bellaïche and Chenevier [**?**] deals, as well, with examples where $d > 1$ in a similar way, achieving interesting results.

5. Articles by Chris Skinner and Eric Urban (See [**?**] and [**?**]) establish the (p-adic) main conjecture for many elliptic curves (defined over $\mathbf{Q}$). Here the reductive group in question is $G = GU(2,2)$, i.e., the general unitary group of signature $(2,2)$ over a quadratic imaginary field $\mathcal{K}/\mathbf{Q}$. The parabolic H is a maximal parabolic subgroup of G fixing an isotropic line; its Levi component is $H = GU(1,1) \times Res_{\mathcal{K}/Q}\mathbf{G}_m$. The "$\pi_o$" is an Eisenstein series induced from the base change to $\mathcal{K}$ of a cuspform on GL_2 over $\mathbf{Q}$ times a Hecke character on $\mathcal{K}$. The p-adically varying family of Galois representations "ρ_t" is a three-variable family corresponding to a "Hida family" times a two-dimensional space of p-adic (degree one) Galois characters over $\mathcal{K}$.

   In [**?**] Skinner and Urban deal with the case where one has a *double zero* for $L(\rho, s)$ at $s = 0$ (when it is the center of the functional equation) to construct *two* linearfly independent extensions of $\mathbf{Q}_p(-1)$ by ρ with appropriate Selmer conditions. See, specifically, loc. cit., Theorem B on page 475. The strategy is to seek generically irreducible deformations of a representation which has (using the terminology of the beginning of this section) *five* irreducible constituents $\rho^{(1)} = \rho$; $\rho^{(2)}$ and $\rho^{(3)}$ trivial; $\rho^{(4)}$ and $\rho^{(4)}$ given by $\mathbf{Q}_p(-1)$.

19 Appendix: Semisimplicity of inertial action in the preferred residual representation mod $p = 691$ of—e.g.—Δ

We will be dealing exclusively with the prime $p = 691$ in this Appendix; for ease of reading though, let p denote the prime 691 below. Let $\Lambda := \mathbf{Z}_p[[\Gamma]]$ where, as usual, $\Gamma \subset \mathbf{Z}_p^*$ is the group of 1-units, and let $\mathbf{T}$ be the finite flat Hida-Hecke Λ-algebra (cf. [?], [?], [?]) that acts naturally on $S_k(\Gamma_1(p); \omega^{k-12}; \mathbf{Z}_p)^o$ for all $k \geq 2$. Here the superscript o means the p-ordinary projection. Let

$$s_k : \Lambda \to \mathbf{Z}_p$$

be specialization to weight k, and denote the Λ-module $\mathbf{Z}_p$ where the Λ-action is given via s_k by the symbol $\mathbf{Z}_p\langle k\rangle$.

Lemma 5. 1. *The natural ring homomorphism*

$$\Lambda \to \mathbf{T}$$

is an isomorphism.

2. *For every k the natural homomorphism of Λ-modules*

$$\mathbf{Z}_p\langle k\rangle \longrightarrow \mathbf{T} \otimes_\Lambda \mathbf{Z}_p\langle k\rangle$$

is an isomorphism.

Proof: Clearly (1) implies (2). Since $\mathbf{T}$ is a finite flat Λ-algebra (1) follows if (2) holds for some value of k. One computes that the $\mathbf{Z}_p$-module $S_{12}(\Gamma_1(p); \omega^0; \mathbf{Z}_p)^o$ is generated by the ordinary eigenform $\Delta(z) - p^{11}\Delta(pz)$. Since $\mathbf{T} \otimes_\Lambda \mathbf{Z}_p\langle 12\rangle$ acts faithfully on $(S_{12}(\Gamma_1(p); \omega^0; \mathbf{Z}_p)^o$ as follows from Hida's theory (cf. Thm. 22 of [?]) we see that (2) holds for $k = 12$, which is enough to prove the lemma. Independently, computations of Stein and Citro establish (2) for $k = 2$.

Hida's theory (cf. loc. cit.) gives a degree two pseudo-representation over $\mathbf{T} = \Lambda$ which associates to the Frobenius element at a prime $\ell \neq 691$ the Hecke operator $T_\ell \in \mathbf{T}$. When specialized to weight two this (pseudo)-representation yields the Galois representation over $\mathcal{F}_v = \mathbf{Q}_{691}$ associated to A_f. Choose the right lattice so that the associated residual representation is $\bar{\rho}$. We can then realize the entire pseudo-representation over $\mathbf{T}$ by an honest Galois representation, $\rho_{\mathbf{T}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{T}) = \mathrm{GL}_2(\Lambda)$ using Proposition 1.6.1 of [?], since the pseudo-representation is residually multiplicity free and $\mathbf{T}$ is a factorial local ring.

The relevant Leopoldt-Kubota L-function $\mathcal{L} \in \Lambda$ has a single zero (and it is of multiplicity 1), or equivalently, $\Lambda/\mathcal{L}\Lambda \simeq \mathbf{Z}_p$. Reducing $\rho_{\mathbf{T}}$ modulo $\mathcal{L}$ one gets a reducible representation

$$\rho_{\mathbf{T},\mathrm{mod}\ \mathcal{L}} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_p)$$

and we choose a basis $\Lambda \oplus \Lambda$ for the underlying Λ-module of the representation $G_{\mathbf{Q}} \to \mathrm{GL}_2(\Lambda)$ so that the action of $G_{\mathbf{Q}}$ is upper-triangular modulo $\mathcal{L}\Lambda$.

In particular, the "sublattice" $M' := \mathcal{L}\Lambda \oplus \Lambda \ \subset \ M := \Lambda \oplus \Lambda$ is $G_{\mathbf{Q}}$-stable so that we have the option of taking either M or M' as our basic lattice in terms of which we will write the representation $\rho_{\mathbf{T}}$. This boils down to considering either the initial representation $\rho_{\mathbf{T}}$ or its conjugate:

$$\begin{pmatrix} \mathcal{L} & 0 \\ 0 & 1 \end{pmatrix} \cdot \rho_{\mathbf{T}} \cdot \begin{pmatrix} \mathcal{L}^{-1} & 0 \\ 0 & 1 \end{pmatrix}.$$

The residual representations attached to these Λ-lattices M and M' (i.e., their reduction modulo the maximal ideal in Λ) are indecomposable representations $G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_p)$ with characters $\mathbf{1}$ and ω^{11} occurring in the two different possible orderings. It follows then that *every* overconvergent eigenform f_k classified by the Λ-representation $\rho_{\mathbf{T}}$ has these—and therefore has the same—two possible indecomposable residual representations.

Definition 3. *Call the indecomposable residual representation for which the character* $\mathbf{1}$ *occurs as a quotient representation, rather than a sub-representation of the residual representation the* **preferred indecomposable residual representation.**

For various reasons—e.g., because it is true for the abelian variety A_f, or—by a simpler route to achieve the same conclusion by consideration of the entire Λ-family (see the second bullet in section **??**)—one learns that the *preferred* residual representation has the property that its inertial action is semisimple. We see, therefore, that the preferred indecomposable residual representation of *any* of the eigenforms classified by this Hida component—has the property that its inertial action is semisimple.

References

[1] J. Bellaïche, Relèvement des formes modulaires de Picard. (French) [Lifting of Picard modular forms] J. London Math. Soc. (2), no. 1, **74** 13–25 (2006)

[2] J. Bellaïche, Congruences endoscopiques et représentations galoisiennes, Thesis, Université Paris-Sud, 2002.

[3] J. Bellaïche, A propos d'un lemme de Ribet, Rend. Sem. Univ. Padova **109** 45-62 (2003)

[4] J. Bellaïche, G. Chenevier, Formes non tempérées pour $U(3)$ et conjectures de Bloch-Kato, Annales scientifiques de l'E.N.S. **37** no. 4, 611-662 (2004)

[5] J. Bellaïche, G. Chenevier, *Families of Galois representations and Selmer groups*, Astérisque **324** (2009)

[6] J. Bernoulli, The Art of Conjecturing, together with Letter to a Friend on Sets in Court Tennis; English Translation and Commentary by Edith Dudley Sylla, The Johns Hopkins Press (2006)

[7] R. Brauer, C. Nesbitt, On the modular characters of groups. Ann. of Math. (2) **42** 556-590 (1941)

[8] J. Buhler, R. Crandall, R. Ernval, T. Metsänkylä, M. Amin Shokrollahie, Irregular primes and cyclotomic invariants to 12 million, Journal of Symbolic Computation, **31** 89-96 (2001)

[9] Cebotarev, N.: Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören, Math. Ann. **95** 191-228 (1926)

[10] G. Chenevier, Familles p-adiques de formes automorphes pour $\mathrm{GL}(n)$, Journal für die reine und angewandte Mathematik **570**, 143-217 (2004)

[11] G. Chenevier, Une correspondance de Jacquet-Langlands p-adique, Duke Math. Journal **126** no. 1, 161-194 (2005)

[12] G. Chenevier, The p-adic analytic space of pseudo-characters of a profinite group and pseudo-representations over arbitrary rings, arXiv:0809.0415v1 [math.NT] Sept. 2 (2008)

[13] L. Clozel On Ribet's Level-raising Theorem for $U(3)$, American Journal of Math, **122** 1265-1287 (2000)

[14] P. Deligne, Formes modulaires et représentations ℓ-adiques, Séminaire Bourbaki, Lect. Notes in Math. **1799** Springer, 139-172 (1971)

[15] Deligne, P.; Rapoport, M. Les schémas de modules de courbes elliptiques. (French) Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 143-316 in Lecture Notes in Math, **349**, Springer (1973)

[16] C. Curtis, I. Reiner, *Representation theory of finite groups and associative algebras,* Wiley (1962)

[17] L. R. Graham, J.M. Kantor, *Naming Infinity* Belknap Press, (2009)

[18] H. Hida, On p-adic Hecke algebras for GL_2 over totally real fields, Ann. of Math. **128** 295-384 (1988)

[19] H. Hida, On nearly ordinary Hecke algebras for GL(2) over totally real fields, Advanced Studies in Pure Math. **17**, 139-169 (1989)

[20] H. Hida, p-Adic ordinary Hecke algebras for GL(2), Ann. de l'Insitut Fourier **44** 1289-1322 (1994)

[21] H. Hida, Control Theorems and Applications, Lectures at Tata institute of fundamental research (Version of 2/15/00) [See http://www.math.ucla.edu/ hida/]

[22] H. Hida, *Hilbert Modular Forms and Iwasawa Theory*, Oxford University Press (2006)

[23] H. Hida, J. Tilouine and E. Urban, Adjoint modular Galois representations and their Selmer groups, Proc. Natl. Acad. Sci. USA **94** , 11121-11124 (1997)

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

[25] M. Kisin, Overconvergent modular forms and the Fontaine-Mazur conjecture, Invent. Math. **153** (2) 373-454 (2003)

[26] S. Lang *Algebraic Number Theory* (Second Edition) Springer (1993)

[27] B. Mazur, A. Wiles, The class field of abelian extensions of $\mathbf{Q}$, Invent. Math. **76** no.2, 179-330 (1984)

[28] I. Piatetski-Shapiro, Two Conjectures on L-functions, pp. 519-522 in *Wolf Prize in Mathematics* Vol 2. (Eds. S.S. Chern and F. Hirzebruch) World Scientific (2000)

[29] M.Raynaud, Schémas en groupes de type $(p, p, \ldots, p)$, Bull. Soc. Math. France. **102** 241-280 (1974)

[30] K. Ribet, A modular construction of unramified p-extensions of $\mathbf{Q}(\mu_p)$, Inventiones math. **34** 151-162 (1976)

[31] D. Rohrlich, Modular Curves, Hecke Correspondences, and L-functions. pp.41-99 in *Modular Forms and Fermat's Last Theorem,* Springer (1997)

[32] J.-P. Serre, J.-P. Serre, *Zeta and L-functions,* Arithmetical Algebraic Geometry, Harper and Row, New York, (1965)

[33] J.-P. Serre, *Abelian ℓ-adic Representations and Elliptic Curves*, W.A., Benjamin Inc. (1968)

[34] J.-P. Serre, Représentations ℓ-adiques, pp. 384-401 in Jean-Pierre Serre/ Oeuvres/ Collected Papers Volume III (1972-1984) Springer (1986)

[35] J.-P. Serre, Sur les représentations modulaires de degré 2 de Gal($\bar{\mathbf{Q}}/\mathbf{Q}$), Duke Math. J. **54** 179-230 (1987)

[36] C. Skinner, Elliptic Curves and Main Conjectures, KUWAIT FOUNDATION LECTURE 49 - May 24, 2005 http://www.dpmms.cam.ac.uk/Seminars/Kuwait/abstracts/L49.pdf

[37] C. Skinner, A. Wiles, Residually reducible representations and modular forms, Journal Publications Mathématiques de L'IHÈS **89**, 6-126 (1999)

[38] C. Skinner, E. Urban, Sur les déformations p-adiques de certaines représentations automorphes. [On the p-adic deformations of certain automorphic representations] J. Inst. Math. Jussieu **5** no. 4. 629-698 (2006)

[39] C. Skinner, E. Urban, Vanishing of L-functions and ranks of Selmer groups, pp. 473-500 in *Proceedings of the International Congress of Mathematicians.* Vol. II, Eur. Math. Soc., Zurich (2006); See also: http://www.math.columbia.edu/ urban/EURP08.html

[40] C. Skinner, E. Urban, The main conjecture for GL_2, See: http://www.math.columbia.edu/ urban/EURP08.html

[41] E. Urban, On residually reducible representations on local rings, J. Algebra **212** no. 2, 738-742 (1999) .

[42] E. Urban, Selmer groups and the Eisenstein-Klingen ideal, Duke Math. J. **106** no. 3, 485-525 (2001) .

[43] E. Urban, Groupes de Selmer et Fonctions L p-adiques pour les représentations modulaires adjointes, See: http://www.math.columbia.edu/ urban/EURP08.html

[44] H. Vandiver, Fermat's Last Theorem: Its history and tHE nature of the known results concerning it, Amer. Math. Monthly, **53** 555-578 (1946); **60** 164-167 (1953)

[45] L. Washington, *Introduction to Cyclotomic Fields*, Springer (1982)

[46] A. Wiles, The Iwasawa conjecture for totally real fields, Ann. of Math. **131**(2), 493 540 (1990)