RECIPROCITY IN THE LANGLANDS PROGRAM SINCE FERMAT'S LAST THEOREM

FRANK CALEGARI

Abstract. The *reciprocity conjecture* in the Langlands program links motives to automorphic forms. The proof of Fermat's Last Theorem by Wiles [180, 169] introduced new tools to study reciprocity. This survey reports on developments using these ideas (and their generalizations) in the last three decades.

1. Introduction

The *reciprocity conjecture* in the Langlands program predicts a relationship between pure motives[1] and automorphic representations. The simplest version (as formulated by Clozel [48, Conj 2.1]) states that there should be a bijection between irreducible motives M over a number field F with coefficients in $\overline{\mathbf{Q}}$ and cuspidal algebraic representations π of $\mathrm{GL}_n(\mathbf{A}_F)$ satisfying a number of explicit additional compatibilities, including the equality of algebraic and analytic L-functions $L(M,s) = L(\pi,s)$. In light of multiplicity one theorems [105], this pins down the correspondence uniquely. There is also a version of this conjecture for more general reductive groups, although its formulation requires some care (as was done by Buzzard and Gee [32]). Beyond the spectacular application by Wiles to Fermat's Last Theorem [180, Theorem 0.5], the Taylor–Wiles method [180, 169] gave a completely new technique — and to this date the most successful one — for studying the problem of reciprocity. The ideas in these two papers have sustained progress in the field for almost[2] 30 years. In this survey, we explain how the Taylor–Wiles method has evolved over this period and where it stands today. **One warning:** the intended audience for this document is entirely complementary to the audience for my talk — I shall assume more than a passing familiarity with the arguments of [180, 169]. Moreover, this survey is as much a personal and historical[3] discussion as a mathematical one — giving anything more than hints on even a fraction of what is discussed here would be close to impossible given

The author was supported in part by NSF Grant DMS-2001097.

[1]Here (in light of the standard conjectures [124]) one may take pure motives up to numerical or homological equivalence. Conjecturally, one can also substitute (for irreducible motive) the notion of an irreducible weakly compatible system of Galois representations [167] or an irreducible geometric Galois representation in the sense of Fontaine–Mazur [83].

[2]Wiles in [180] dates the completion of the proof to September 19, 1994.

[3]A whiggish history, naturally. Even with this caveat, it should be clear that the narrative arc of progress presented here at best represents my own interpretation of events. I have added a few quotes from first hand sources when I felt they conveyed a sense of what the experts were thinking in a manner not easily obtainable from other sources. For other survey articles on similar topics, see [30, 24].

the space constraints and the competence of the author. Even with the absence of any real mathematical details in this paper, the sheer amount of activity in this field has led me to discard any discussion of advances not directly related to $R = \mathbf{T}$ theorems, which necessitates the omission of a lot of closely related beautiful mathematics.

1.1. **The Fontaine–Mazur Conjecture.** Let F be a number field. The Fontaine–Mazur conjecture[4] [83] predicts that any continuous irreducible p-adic Galois representation

$$\rho : G_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p)$$

which is both unramified outside finitely many primes and potentially semi-stable (equivalently, de Rham [56]) at all places $v|p$ should be associated to a motive M/F with coefficients in $\overline{\mathbf{Q}}$. Any such ρ is automatically conjugate to a representation in $\mathrm{GL}_n(E)$ for some finite field $E/\mathbf{Q}_p$ and further stabilizes an $\mathcal{O}_E$-lattice. The corresponding residual representation $\overline{\rho} : G_F \to \mathrm{GL}_n(k)$ where $k = \mathcal{O}_E/\pi_E$ is the residue field of E is unique up to semi-simplification. Let us assume here for expositional convenience that $\overline{\rho}$ is absolutely irreducible. Following Mazur [129], one may define a universal deformation ring which parameterizes all deformations of $\overline{\rho}$ unramified outside a finite set S. One can then further impose local conditions to define deformation rings R whose $\overline{\mathbf{Q}}_p$-valued points are associated to Galois representations which are de Rham at $v|p$ with fixed Hodge–Tate weights. Assuming the Fontaine–Mazur conjecture, these $\overline{\mathbf{Q}}_p$-valued points correspond to all pure motives M unramified outside S whose p-adic realizations are Galois representations with the same local conditions at p and the same fixed residual representation $\overline{\rho}$. Assuming the reciprocity conjecture, these motives should then be associated to a finite dimensional space of automorphic forms. This leads to the extremely non-trivial prediction that R has finitely many $\overline{\mathbf{Q}}_p$-valued points. The problem of reciprocity is now to link these $\overline{\mathbf{Q}}_p$-valued points of R to automorphic forms.

1.2. $R = \mathbf{T}$ **theorems.** Associated to the (conjectural) space of automorphic forms corresponding to $\overline{\mathbf{Q}}_p$-valued points of R is a ring of endomorphisms generated by Hecke operators. The naïve version of $\mathbf{T}$ is defined to be the completion of this ring with respect to a maximal ideal $\mathfrak{m}$ defined in terms of $\overline{\rho}$. The mere existence of $\mathfrak{m}$ is itself conjectural, and amounts — in the special case of odd absolutely irreducible 2-dimensional representations $\overline{\rho}$ of $G_{\mathbf{Q}}$ — to Serre's conjecture [156]. Hence, in the Taylor–Wiles method, one usually assumes the existence of a suitable $\mathfrak{m}$ as a hypothesis. The usual shorthand way of describing what comes out of the Taylor–Wiles method is then an "$R = \mathbf{T}$ theorem." Proving an $R = \mathbf{T}$ theorem can more or less be divided into three different problems:

[4] Fontaine told me (over a *salad de gésiers* in Roscoff in 2009) that he and Mazur formulated their conjecture in the mid-80s. (Colmez pointed me towards these notes [81] from a talk given by Fontaine at the 1988 Mathematische Arbeitstagung in Bonn.) He noted that Serre had originally been skeptical, particularly of the claim that any everywhere unramified representation inside $\mathrm{GL}_n(\overline{\mathbf{Q}}_p)$ must have finite image, and set off to find a counterexample (using the construction of Golod–Shavarevich [91]). He (Serre) did not succeed!

(1) Understanding $\mathbf{T}$. Why does there exist[5] a map $R \to \mathbf{T}$? This is the problem of the "existence of Galois representations." Implicit here is the problem of showing that those Galois representations not only exist but have the "right local properties" at the ramified primes, particularly those dividing p.
(2) Understanding R. Wiles introduced a mechanism for controlling R via its tangent space using Galois cohomology (in particular Poitou–Tate duality [131]), and this idea has proved remarkably versatile. What has changed, however, is our understanding of local Galois representations and how this information can be leveraged to understand the structure of R.
(3) Understanding why the map $R \to \mathbf{T}$ is an isomorphism.

We begin by summarizing the original $R = \mathbf{T}$ theorem from this viewpoint (or more precisely, the modification by Faltings which appears as an appendix to [169]). We only discuss for now the so-called "minimal case[6]" since this is most relevant for subsequent generalizations (see §6.2). Our summary is cursory, but see [68, 67] for excellent expositional sources on early versions of the Taylor–Wiles method. We start with a representation $\overline{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ for $p > 2$ which (say) comes from a semistable elliptic curve E and which we assume to be modular. By a theorem of Ribet [149], we may assume it is modular of level either $N = N(\overline{\rho})$ or $N = N(\overline{\rho})p$ where $N(\overline{\rho})$ is the Serre weight [156] of $\overline{\rho}$.

(1) Understanding $\mathbf{T}$: The construction of Galois representations associated to modular forms has its own interesting history (omitted here), but (in the form originally needed by Wiles) was more or less complete for modular forms (and even Hilbert modular forms) by 1990. The required local properties at primes different from p followed from work of Carayol [47], and the local properties at p were well understood either by Fontaine–Laffaille theory [82], or, in the ordinary case, by Mazur–Wiles [130] (see also work of Hida [101, 102]).
(2) Understanding R: Here R is a deformation ring of $\overline{\rho}$ subject to precise local deformation conditions at p and the primes dividing $N(\overline{\rho})$. For the prime p, the local conditions amount either to an "ordinary" or "finite-flat" restriction. One then interprets the dual of the reduced tangent space $\mathfrak{m}_R/(\mathfrak{m}_R^2, p)$ of R in terms of Galois cohomology, in particular as a subgroup (Selmer group) of classes in $H^1(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$ satisfying local conditions. This can be thought of as analogous to a class group, and one does not have any *a priori* understanding of how large it can be although it has some finite dimension d. Using the Greenberg–Wiles formula, the obstructions in $H^2(\mathbf{Q}, \mathrm{ad}^0(\overline{\rho}))$ can be related to the reduced tangent space, and allow one to realize R as a quotient of $W(k)[\![x_1, \ldots, x_d]\!]$ by d relations. In particular, if R was finite and free as a $W(k)$-module (as would be the case if $R = \mathbf{T}$) then R would be a complete intersection.

[5] At the time of Wiles' result, this was seen as the easier direction (if not easy), although, in light of the success of the Taylor–Wiles method, it may well be the harder direction in general.

[6] The case when the Galois representations attached to R and $\mathbf{T}$ have minimal level N as determined by the residual representation.

(3) Understanding why the map $R \to \mathbf{T}$ is an isomorphism. Here lies the heart of the Taylor–Wiles method. The ring $\mathbf{T}$ acts on a natural module M of modular forms. One shows — under a mild hypothesis on $\overline{\rho}$ — the existence of (infinitely many) sets $Q = Q_N$ for any natural number N of cardinality $|Q| = d$ — so-called Taylor–Wiles primes — with a number of pleasant properties:
   (a) The primes $q \in Q$ are congruent to 1 mod p^N.
   (b) Let R_Q be the deformation ring capturing the same local properties as R but modified so that the representations at primes in Q may now be ramified of degree p^N. There is naturally a surjection $R_Q \to R$, but for Taylor–Wiles primes, this modification does not increase the size of the tangent space. In particular, for a fixed ring $R_\infty = W(k)[\![x_1, \ldots, x_d]\!]$ there are surjections $R_\infty \to R_Q \to R$ for every Q.
   (c) The corresponding rings $\mathbf{T}$ and $\mathbf{T}_Q$ act naturally on spaces of modular forms M and M_Q respectively. Using multiplicity one theorems, Wiles proves (see [180, Theorem 2.1]) that M and M_Q are free of rank one over $\mathbf{T}$ and $\mathbf{T}_Q$ respectively. The space M can be interpreted as a space of modular forms for a particular modular curve X. The second key property of Taylor–Wiles primes is that there are no new modular forms associated to $\overline{\rho}$ at level $X_0(Q)$, and hence M can *also* be interpreted as a space of modular forms for $X_0(Q)$. There is a Galois cover $X_1(Q) \to X_0(Q)$ with Galois group $(\mathbf{Z}/Q\mathbf{Z})^\times$, and hence an intermediate cover $X_H(Q) \to X_0(Q)$ with Galois group $\Delta_N = (\mathbf{Z}/p^N\mathbf{Z})^d$ acting via diamond operators. The space M_Q is essentially a localization of a certain space of modular forms for $X_H(Q)$ (with some care taken at the Hecke operators for primes dividing Q). Since the cohomology of modular curves (localized at the maximal ideal corresponding to $\mathfrak{m}$) is concentrated in degree one, the module M_Q turns out to be free over an auxiliary ring $S_N = W(k)[\Delta_N]$ of diamond operators, and the quotient $M_Q/\mathfrak{a}_Q$ for the augmentation ideal $\mathfrak{a}_Q$ of S_N is isomorphic to M. It follows that $\mathbf{T}_Q/\mathfrak{a}_Q = \mathbf{T}$.
   (d) The diamond operators have an interpretation on the Galois deformation side, and there is a identification $R_Q/\mathfrak{a}_Q = R$ where R_Q and $\mathbf{T}_Q$ can be viewed compatibly as S_N-modules.

(4) Finally, one "patches" these constructions together for larger and larger Q. This is somewhat counterintuitive, since for different Q the Galois representations involved are not compatible. However, one forgets the Galois representations and only remembers the structures relative to both the diamond operators S_N and R_∞, giving the data of a surjection

$$R_\infty \to \mathbf{T}_\infty$$

with a compatible action of $S_\infty = \operatorname{proj\,lim} S_N \simeq W(k)[\![t_1, \ldots, t_d]\!]$. Using the fact that $\mathbf{T}_\infty$ is free of finite rank over S_∞, and that R_∞ and S_∞ are formally smooth of the same dimension, one deduces that $R_\infty = \mathbf{T}_\infty$ and then $R = \mathbf{T}$ after quotienting out by the augmentation ideal of S_∞.

2. The early years

2.1. **The work of Diamond and Fujiwara.** Wiles made essential uses of multiplicity one theorems in order to deduce that M_Q was free over $\mathbf{T}_Q$. Diamond [72] and Fujiwara [85] (independently) had the key insight that one could instead patch the modules M_Q directly — and then argue directly with the resulting object M_∞ as a module over R_∞ which was also free over S_∞. Using the fact that R_∞ is formally smooth, this allowed one to deduce *a posteriori* that M_∞ was free over R_∞ using the Auslander–Buchsbaum formula [9]. This not only removed the necessity of proving difficult multiplicity one results but gave new proofs of these results[7] which could then be generalized to situations where the known methods (often using the q-expansion principle) were unavailable[8]. Diamond had the following to say about how he came up with the idea to patch modules rather than use multiplicity one theorems:

> My vague memory is that I was writing down examples of ring homomorphisms and modules, subject to some constraints imposed by a Taylor–Wiles setup, and I couldn't break "M free over the group ring implies M free over R." (I still have the notebook with the calculations somewhere, mostly done during a short trip with some friends to Portugal.) I didn't know *what* commutative algebra statement I needed, but I knew I needed to learn more commutative algebra and found my way to Bruns and Herzog's "Cohen-Macaulay Rings" [28] (back in the library in Cambridge UK by then). When I saw the statement of Auslander–Buchsbaum, it just clicked.

Diamond made a second improvement [70, 71] dealing with primes away from p in situations where the corresponding minimal local deformation problem was not controlled by the Serre level $N(\overline{\rho})$ alone.

2.2. **Integral p-adic Hodge Theory, part I: Conrad–Diamond–Taylor.** One early goal after Fermat was the resolution of the full Taniyama–Shimura conjecture, namely, the modularity of all elliptic curves over $\mathbf{Q}$. After the improvements of Diamond, the key remaining problem was understanding deformation rings associated to local Galois representations at p coming from elliptic curves with bad reduction at p. Since Wiles' method (via Langlands–Tunnell [127, 177]) was ultimately reliant on working with the prime $p = 3$, this meant understanding deformations at p of level p^2 and level p^3, since any elliptic curve over $\mathbf{Q}$ has a twist such that the largest power of 3 dividing the conductor is at most 27. Ramakrishna in his thesis [145] had studied the local deformation problem for finite flat representations (the case when $(N, p) = 1$) and proved that the corresponding local deformation rings were formally smooth. The case when p exactly divides N was subsumed into the ordinary case, also treated by Wiles. In level p^2, one can show that the Galois representations associated to the relevant modular forms[9] of level p^2 become finite flat after passing to a finite extension $L/\mathbf{Q}_p$ with ramification degree $e \le p - 1$. In

[7] There is an intriguing result of Brochard [27] which weakens the hypotheses of Diamond's freeness criterion even further, although this idea has not yet been fully exploited.

[8] The history of the subject involves difficult theorems in the arithmetic geometry of Shimura varieties being replaced by insights from commutative algebra, paving the way to generalizations where further insights from the arithmetic geometry of Shimura varieties are required.

[9] This is not true for all modular forms of level p^2 and weight 2, but only for those whose conductor at p remains divisible by p^2 after any quadratic twist.

this range, Conrad [64, 65] was able to adapt ideas of Fontaine [80] to give an equivalence between the local Galois deformations (assuming $\overline{\rho}|_{G_{\mathbf{Q}_p}}$ was irreducible) and linear algebra data. In particular, as in the work of Ramakrishna, one can show that the relevant local deformation rings are formally smooth, and so Conrad, Diamond, and Taylor were able to adapt the Taylor–Wiles method to this setting [66].

2.3. **Integral p-adic Hodge Theory, part II: Breuil–Conrad–Diamond–Taylor.** A central technical ingredient in all of the arguments so far has been some use of integral p-adic Hodge Theory, and in particular the theory of finite flat group schemes and Barsotti–Tate groups as developed by Fontaine and others. All integral versions of this theory required a hypothesis on either the weight or the ramification index e relative to the bound $p-1$. However, around this time, Christophe Breuil made a breakthrough[10] by finding a new way to understand the integral theory of finite flat group schemes over arbitrarily ramified bases [19]. This was just the technical tool required to push the methods of [66] to level p^3. Using these results, Breuil, Conrad, Diamond and Taylor [25] were able to show that enough suitably chosen local deformation rings were formally smooth to prove the modularity of all elliptic curves.

2.4. **Higher weights, totally real fields, and base change.** Many of the methods which worked for modular forms were directly adaptable to the case both of general rank 2 motives over $\mathbf{Q}$ with distinct Hodge–Tate weights (corresponding to modular forms of weight $k \geq 2$ rather than $k = 2$) and also to such motives over totally real fields (which are related to Hilbert modular forms), see in particular the work of Fujiwara [85] (and more recently Frietas–Le Hung–Siksek [84]). Another very useful innovation was a base change idea of Skinner–Wiles [161] which circumvented the need to rely on Ribet's level lowering theorem. The use of cyclic base change ([127] in this case and [5] in general) subsequently became a standard tool in the subject. For example, it meant that one could always reduce to a situation where the ramification at all primes $v \nmid p$ was unipotent. The paper [161] was related to a more ambitious plan by Wiles to prove modularity for all totally real fields:

> After Fermat I started to work with Taylor and then Diamond on the general case but decided very soon that I would rather try to do the totally real case for GL(2). I think this was while I was getting back into other kinds of problems but I thought I should still earn my bread and butter. One lunch time at the IAS in 1996 Florian Pop spoke to me and explained to me about finding points over fields totally split at some primes (e.g. real places) as he had written a paper [92] about this with some others. Was this any use for the Tate–Shavarevich group? I immediately saw that whether or not it was any use for TS (I doubted it) it should certainly give potential modularity. This gave some kinds of lifting so I worked on the other half (i.e. descent) thinking that just needed a similar insight. At some point I suggested to Chris that we try to do Ribet's theorem using cyclic base change as that would be useful in proving modularity and was buying time while I waited to get the right idea. Unfortunately I completely misjudged the difficulty of descent and the problem is still there. I think it is both much harder than I thought and also more

[10]Much of the development of integral p-adic Hodge theory over the last 20 years since [25] has been inspired by its use in the Taylor–Wiles method. However, the timing of Breuil's work was more of a happy coincidence, although Breuil was certainly aware of the fact that a computable theory of finite flat group schemes over highly ramified bases could well have implications in the Langlands program.

important. I hope still to prove it! Of course Taylor found potential modularity and then, what I had assumed was much harder, a way to think about $\mathrm{GL}(n)$.

3. Reducible representations: Skinner–Wiles

One of the key hypotheses in the Taylor–Wiles method concerns restrictions on the representation $\overline{\rho}$, in particular the hypothesis that $\overline{\rho}|_{G_{\mathbf{Q}(\zeta_p)}}$ is absolutely irreducible. In [162, 160, 159], Skinner and Wiles introduced a new argument in which this hypothesis was relaxed, at least assuming the representations were ordinary at p. In the ordinary setting, one can replace the rings R and $\mathbf{T}$ (which in the original setting are finite over $W(k)$) by rings which are finite (and typically flat) over Iwasawa algebras $\Lambda = W(k)[\![(\mathbf{Z}_p)^d]\!]$ for some d which arise as weight spaces, the point being that the ordinary deformations of varying weight admit a good integral theory. The first innovation (in part) involves making a base change so that the reducible locus is (relatively) "small," (measured in terms of the codimension over Λ). The second idea is then to apply a variant of the Taylor–Wiles method to representations $\varrho : G_F \to \mathrm{GL}_2(\mathbf{T}/\mathfrak{p})$ for non-maximal prime ideals $\mathfrak{p}$ of F[11]. Wiles again:

> We had worked out a few cases we could do without big Hecke rings in some other papers and I would say it was more a feat of stamina and technique to work through it. Of course the use of these primes was much more general and systematic than anything that went before. There is also an amusing point in this paper where we use a result from commutative algebra. It seemed crucial then though I don't know if it still is. This is proposition A.1 of Raynaud [148]. I had thought at some point during the work on Fermat that this result might be needed and had asked Michel Raynaud about it. He said he would think about it. A week later he came back to me, somewhat embarrassed that he had not known right away, to say that it was a result in his wife's thesis. So the reference to M.Raynaud is actually to his wife, Michèle Raynaud, though he gave the reference.

Allen [2] was later able to adapt these arguments to the $p = 2$ dihedral case, which (in a certain sense) realized the original desire[12] of Wiles to work at the prime $p = 2$.

4. The Artin Conjecture

While the approach of [180, 169] applied (in principle) to all Galois representations associated to modular forms of weight $k \geq 2$, the case of modular forms of weight $k = 1$ is qualitatively quite different (see also §10.1). It was therefore quite surprising when Buzzard–Taylor [33] proved weight one modularity lifting theorems for odd continuous representations $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ which were unramified at p. Using this, Buzzard–Dickinson–Shepherd-Barron–Taylor [31] proved the Artin conjecture for a positive proportion of all odd A_5 representations, which had previously only been known in a finite number

[11] Representations ϱ to infinite quotients $\mathbf{T}/\mathfrak{p}$ had also arisen in Wiles' paper on Galois representations associated to ordinary modular forms [179] where the concept of pseudo-deformation was also first introduced.

[12] As far as primary historical sources go, the introduction of Wiles' paper [180] is certainly worth reading.

of cases[13] up to twist. Standard ordinary modularity theorems showed the existence of ordinary modular forms associated to such representations ρ — however, the classicality theorems of Hida [101] do not apply (and are not true!) in weight one. The main idea of [33] was to exploit the fact that ρ is unramified to construct *two* ordinary modular forms each corresponding to a choice of eigenvalue of $\rho(\mathrm{Frob}_p)$ assuming these eigenvalues are distinct[14]. One then has to argue [33] that these two ordinary forms are the oldforms associated to a classical eigenform of weight one, which one can do by exploiting both the rigid geometry of modular curves and the q-expansion principle.

Although the original version of this argument required a number of improvements to the usual Taylor–Wiles method (Dickinson overcame some technical issues when $p = 2$ [73] and Shepherd-Barron–Taylor proved some new cases of Serre's conjecture for $\mathrm{SL}_2(\mathbf{F}_4)$ and $\mathrm{SL}_2(\mathbf{F}_5)$-representations in [157]), it was ripe for generalization to totally real fields[15] After a key early improvement by Kassaei [106], the $n = 2$ Artin conjecture for totally real fields is now completely resolved under the additional assumption that the representation is odd by a number of authors, including Kassaei–Sasaki–Tian and Pilloni–Stroh [108, 107, 109, 152, 141, 143]. On the other hand, the reliance on q-expansions in this argument has proved an obstruction to extending this to other groups. (See also §11.2).

5. Potential Modularity

One new idea which emerged in Taylor's paper [166] was the concept of *potential modularity*. Starting with a representation $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ for a totally real field F, one could sidestep the (difficult) problem of proving the modularity of $\overline{\rho}$ by proving it was modular over some finite totally real extension F'/F. In the original paper [180], Wiles employed a 3-5 switch to deduce the modularity of certain mod 5 representations from the modularity of mod 3 representations. More generally, one can prove the modularity of a mod p representation $\overline{\rho}_p$ from the modularity of a mod q representation $\overline{\rho}_q$ if one can find both of them occurring as the residual representation of a compatible family where the Taylor–Wiles hypotheses apply to $\overline{\rho}_q$. For example, if $\overline{\rho}_p$ and $\overline{\rho}_q$ are representations valued in $\mathrm{GL}_2(\mathbf{F}_p)$ and $\mathrm{GL}_2(\mathbf{F}_q)$ respectively, one can try to find the compatible family by finding an elliptic curve with a given mod p and mod q representation. The obstruction to doing such a p-q switch over F is that the corresponding moduli spaces (which in this case are twists of the modular curve $X(pq)$) are not in general rational, and hence have no reason to admit rational points. However, exploiting an idea due to Moret-Bailly [132], Taylor showed that these moduli spaces at least had many points over totally real fields where one

[13]In a computational *tour de force* for the time, Buhler [29] in his thesis had previously established the modularity of an explicit odd projective A_5 representation of conductor 800.

[14]This argument can be modified to deal with the case when the eigenvalues of $\overline{\rho}(\mathrm{Frob}_p)$ coincide by modifying R and $\mathbf{T}$ to include operators corresponding (on the Hecke side) to U_p. Geraghty and I discovered an integral version of this idea ourselves ("doubling," following Wiese's paper [178]) during the process of writing [38], although it turned out that, at least in characteristic zero, Taylor already had the idea in his back pocket in the early 2000s.

[15]The proof all that finite odd 2-dimensional representations over $\mathbf{Q}$ are modular was completed by Khare and Wintenberger as a consequence of their proof of Serre's conjecture, see §8.

could additionally ensure that the Taylor–Wiles hypothesis applies at the prime q. At the cost of proving a weaker result, this gives a huge amount of extra flexibility that has proved remarkably useful. Taylor's first application of this idea was to prove the Fontaine–Mazur conjecture for many 2-dimensional representations, since the potential modularity of these representations was enough to prove (for example) that they come from compatible families of Galois representations (even over the original field F!), and that they satisfy purity (which is known for Hilbert modular forms of regular weight). The concept of potential modularity, however, has proved crucial for other applications, not least of which is the proof of the Sato–Tate conjecture (see §9.2).

6. The work of Kisin

A key ingredient in the work of Breuil–Conrad–Diamond–Taylor (§2.2,§2.3) (and subsequent work of Savitt [153, 154]) was the fact that a certain local deformation ring R^{fl} defined in terms of integral p-adic Hodge theory was formally smooth. The calculations of [25, 154], however, applied only to some (very) carefully chosen situations sufficient for elliptic curves but certainly not for all 2-dimensional representations. In the 2000s, Kisin made a number of significant contributions, both to the understanding of local deformation rings but also to the structure of the Taylor–Wiles argument itself [117, 118, 121, 122, 123, 119].

6.1. **Local deformation rings at $v = p$.** One difficulty with understanding local deformation rings R^{fl} associated to finite flat group schemes over highly ramified bases is that the group schemes themselves are not uniquely defined by their generic fibres. Kisin [122] had the idea that one could also define the moduli space of the group schemes themselves, giving a projective resolution $\mathscr{GR} \to \mathrm{Spec}(R^{\mathrm{fl}})$ (this map is an isomorphism after inverting p). Kisin further realized that the geometry of $\mathscr{GR}$ was related to local models of Shimura varieties, for which one had other available techniques to analyze their structure and singularities. Later, Kisin was also able [119] to construct local deformation rings R capturing deformations of a fixed local representation $\overline{\rho}$ which become semi-stable over a fixed extension $L/\mathbf{Q}_p$ and had Hodge–Tate weights in any fixed finite range $[a, b]$, absent a complete integral theory of such representations. (There are also are constructions where one fixes the inertial type of the corresponding representation.) Kisin further proved that the generic fibres of these rings were indeed of the expected dimension and often formally smooth.

6.2. **Kisin's modification of Taylor–Wiles.** Beyond analyzing the local deformation rings themselves, Kisin crucially found a way [122] to modify the Taylor–Wiles method to avoid the requirement that these rings are formally smooth, thus greatly expanding the scope of the method. First of all, Kisin reimagined the global deformation ring R as an algebra over a (completed tensor product)

$$R^{\mathrm{loc}} = \widehat{\bigotimes}_{v \in S} R_v$$

of local deformation rings R_v for sets of places $v \in S$, in particular including the prime p[16]. Now, after a Taylor–Wiles patching argument, one constructs a big module M_∞ over R_∞ (and free over the auxiliary ring of diamond operators S_∞) but where R_∞ is no longer a power series ring over $W(k)$ but a power series ring over R^{loc}. If the algebras R_v for $v \in S$ are themselves power series rings, one is reduced precisely to the original Taylor–Wiles setting as modified by Diamond. On the other hand, if the R_v are (for example) not power series rings but are integral domains over $W(k)$ of the expected dimension, then Kisin explained how one could still deduce that $M[1/p]$ was a faithful $R[1/p]$-module, which proves that $R[1/p] = \mathbf{T}[1/p]$ and suffices for applications to modularity. More generally, assuming only that the R_v are flat over $W(k)$ and that the generic fibre $R_v[1/p]$ is equidimensional of the expected dimension, the modularity of any point of R reduces to showing that there is at least one modular point which lies on the same component of $R_v[1/p]$[17].

In the original modularity lifting arguments, one treated the minimal case first and then deduced the non-minimal cases using a subtle commutative algebra criterion which detected isomorphisms between complete intersections. From the perspective of Kisin's modification, all that is required is to show that there exists a single modular point with the right non-minimal local properties. In either case, both Wiles and Kisin used Ihara's Lemma to establish the existence of congruences between old and new forms, but Kisin's argument is much softer and thus more generalizable to other situations[18]. Kisin had the following to say about his thought process:

> The idea of thinking of R as an R^{loc} algebra just popped into my head, after I'd been thinking about the Wiles–Poitou–Tate formula, and how it fit into the Taylor–Wiles patching argument. This was in Germany, I think in 2002. I had the idea about moduli of finite flat group schemes in the Fall of 2003, after I arrived in Chicago. It was entirely motivated by modularity. I had been trying to compute these deformation rings, by looking at deformations of finite flat group schemes. For $e < p - 1$, the finite flat model is unique, so I knew this gave the deformation ring in this case; this already gave some new cases. However I was stuck about the meaning of these calculations in general for quite some time. At some point I thought I'd better write up what I had, but as soon as I started thinking about that — within a day — I realized what the correct picture was with the families of finite flat group schemes resolving the deformation ring. I already knew about Breuil's unpublished note [18], and quite quickly was able to prove the picture was correct. It was remarkable that prior to coming to Chicago, I didn't even know the

[16] Since the local residual representations are typically reducible, Kisin also introduced the notion of *framed* deformation rings which are always well-defined, and which (properly taking into account the extra variables) are compatible with the Taylor–Wiles argument.

[17] There are some subtleties to understanding $R[1/p]$ for complete local Noetherian $W(k)$-algebras that are not obvious on first consideration. The first and most obvious blunder to avoid is to recognize that $R[1/p]$ is usually far from being a local ring. Similarly, the ring $R[1/p]$ can be regular and still have multiple components, as can be seen in an example as simple as $R = \mathbf{Z}_p[\![X]\!]/X(X-p)$. Perhaps more importantly, however, the ring $R[1/p]$ "behaves" in some important ways like a finitely generated algebra over a field.

[18] In particular, Wiles' numerical criterion [68, Thm 5.3] relies on certain rings being complete intersections, and Kisin's local deformation rings are not complete intersections (or even Gorenstein) in general — see [163].

definition of the affine Grassmannian, but within a few months of arriving, it actually showed up in my own work.

To me the whole project was incredibly instructive. If I had known more about what was (thought to be) essential in the Taylor–Wiles method, I never would have started the project. Not having fixed ideas gave me time to build up intuition. I also should have gotten the idea about moduli of finite flat group schemes much sooner if I'd been more attentive to what the geometry was trying to tell me.

7. p-adic local Langlands

7.1. **The Breuil–Mézard conjecture.** Prior to Kisin's work, Breuil and Mézard [26] undertook a study of certain low weight potentially semi-stable deformation rings, motivated by [25]. They discovered (in part conjecturally) a crucial link between the geometry of these Galois deformation rings (in particular, the Hilbert–Samuel multiplicities of their special fibres) with the mod-p reductions (and corresponding irreducible constituents) of lattices inside locally algebraic p-adic representations of $\mathrm{GL}_2(\mathbf{Z}_p)$. In the subsequent papers [20, 21], Breuil raised the hope that there could exist a p-adic Langlands correspondence relating certain mod-p (or p-adic Banach space) representations of $\mathrm{GL}_2(\mathbf{Q}_p)$ to geometric 2-dimensional p-adic representations of $G_{\mathbf{Q}_p}$[19]. Breuil recounts the origins of these conjectures as follows:

> The precise moment I became 100% sure that there would be a non-trivial p-adic correspondence for $\mathrm{GL}_2(\mathbf{Q}_p)$ was in the computations of [21]. In these computations, I reduced mod p certain $\overline{\mathbf{Z}}_p$–lattices in certain locally algebraic representations of $\mathrm{GL}_2(\mathbf{Q}_p)$, and at some point, I found out that this reduction mod p had a really nice behaviour, so nice that clearly, it was predicting (via the mod-p correspondence) what the reduction mod-p would be on the $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$-side.

These ideas were further developed by Colmez in [58, 59, 60] amongst other papers[20]: Colmez studied various Banach space completions defined by Breuil and proved they were non-zero using the theory of (φ,Γ)-modules. Since the theory of (φ,Γ)-modules applies to all Galois representations and not just potentially semi-stable ones, this led Colmez to propose a p-adic local Langlands correspondence for *arbitrary* 2-dimensional representations $G_{\mathbf{Q}_p} \to \mathrm{GL}_2(E)$, and he was ultimately able to construct a functor from suitable $\mathrm{GL}_2(\mathbf{Q}_p)$-representations to Galois representations of $G_{\mathbf{Q}_p}$. Colmez gave a talk on his construction at a conference in Montreal in September 2005. At the same conference, Kisin gave a talk presenting a proof of the Breuil–Mézard conjecture by relating it directly

[19] The starting observation [22] is as follows: if $\pi = \bigotimes' \pi_v$ is the automorphic representation associated to a modular form f, then π_v determines (and is determined by) $\rho_f|_{G_{\mathbf{Q}_v}}$ for all $v \neq p$ (at least up to Frobenius semi-simplification). On the other hand, π_p does not determine the p-adic representation $\rho_f|_{G_{\mathbf{Q}_p}}$ (except in the exceptional setting where π_p is spherical and a_p is not a p-adic unit), raising the question of what extra $\mathrm{GL}_2(\mathbf{Q}_p)$ structure associated to f should determine (and be determined by) $\rho_f|_{G_{\mathbf{Q}_p}}$.

[20] In [116], Kisin had shown that the p-adic representations V associated to non-classical finite slope overconvergent modular forms with U_p-eigenvalue a_p satisfied $\dim D_{\mathrm{cris}}(V) = 1$, and moreover that crystalline Frobenius acted on this space by a_p. (This paper was itself apparently motived by the goal of *disproving* the Fontaine–Mazur conjecture!) On the way to the 2004 Durham symposia on L-functions and Galois representations, Fontaine raised the question to Colmez to what extent this determined the corresponding Galois representation. Colmez worked out the answer the evening before his talk and incorporated it into his lecture the following day, ultimately leading to the notion of trianguline representations [58].

to $R = \mathbf{T}$ theorems and the Fontaine–Mazur conjecture for odd 2-dimensional representations of $G_{\mathbf{Q}}$ with distinct Hodge–Tate weights. While Kisin's argument exploited results of Berger–Breuil [14] and Colmez, it was realized by the key participants (perhaps in real time) that Colmez' p-adic local Langlands correspondence should be viewed as taking place over the entire local deformation ring. Subsequently Colmez was able to construct the inverse functor[21]. Colmez writes:

> I received a paper of Breuil (a former version of [23]) during my stay at the Tata Institute in December 2003—January 2004. In December, I was spending Christmas under Goa's palm trees with my daughter when Breuil's paper arrived in my email. That paper contained a conjecture (in the semi-stable case) that I was sure I could prove using (φ, Γ)-modules (if it was true...). I spent January 2004 working on it and after 15 days of computations in the dark, I finally found a meaning to some part of a painful formula (you can find some shadow of all of this in (iii) of Remark 0.5 of my unpublished [57]). By the end of the month, I was confident that the conjecture was proved and I told so to Breuil who adapted the computations to the crystalline case, and wrote them down with the help of Berger (which developed into [14]). (One thing that makes computations easier and more conceptual in the crystalline case is that you end up with the universal completion of the locally algebraic representation you start with; something that is crucial in Matthew [Emerton]'s proof of the FM conjecture.) Durham was in August of that year and Berger–Breuil had notes from a course they had given in China [13]. Those notes were instrumental in my dealing with trianguline representations at Durham (actually, I did some small computation and the theory just developed by itself during the night before my talk which was supposed to be on something else...I think I came up with the concept of trianguline representations later, to justify the computations, I don't remember what language I used in my talk which had some part on Banach–Colmez spaces as far as I can remember.

7.2. **Local–global compatibility for completed cohomology.** From a different perspective, Emerton had introduced the completed cohomology groups [77] as an alternative means for constructing the Coleman–Mazur eigencurve [54]. Inspired by Breuil's work, Emerton formulated [76] a local–global compatibility conjecture for completed cohomology in the language of the then nascent p-adic Langlands correspondence. After the construction of the correspondence for $\mathrm{GL}_2(\mathbf{Q}_p)$ by Colmez and Kisin, Emerton was able to prove most of his conjecture, leading to a new proof of (many cases of) the Fontaine–Mazur conjecture. The results of Kisin [120] and Emerton fell short of proving the full version of this conjecture for two reasons. The first was related to some technical issues with the p-adic local Langlands correspondence, both at the primes $p = 2$ and 3 but also when the residual representation locally had the shape $1 \oplus \overline{\varepsilon}$ for the cyclotomic character ε. (The local issues have now more or less all been resolved [63]. The most general global

[21]To add some further confusion to the historical chain of events, the published version of [120] incorporates some of these subsequent developments. Note also that the current state of affairs is that the proof of the full p-adic local Langlands correspondence for $\mathrm{GL}_2(\mathbf{Q}_p)$ (for example as proved in [63] but see also [60, Remarque VI.6.51]) still relies on the global methods of [78], which in turn relies on [60]. These mutual dependencies, however, are not circular! The difficulty arises in the supercuspidal case. One philosophical reason that global methods are useful here is that all global representations are yoked together by an object (the completed cohomology group $\widetilde{H}^1(\mathbf{Z}_p)$) with good finiteness properties. One can then exploit the fact that crystabeline representations (for which the p-adic local Langlands correspondence is known by [60]) are Zariski dense inside unrestricted global deformation rings ([78, Theorem 1.2.3], using arguments going back to Böckle [15]).

results for $p = 2$ are currently due to Tung [176].) A second restriction was the Taylor–Wiles hypothesis that $\overline{\rho}$ was irreducible. Over the intervening years, a number of key improvements to the local story have been found, in particular by Colmez, Dospinescu, Hu, and Paškūnas [140, 63, 104]. Very recently, Lue Pan [137] found a way to marry techniques from Skinner–Wiles in the reducible case (§3) to techniques from p-adic local Langlands to completely prove the modularity (up to twist) of any geometric representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ for $p \geq 5$ only assuming the hypotheses that ρ has distinct Hodge–Tate weights and that ρ is odd[22].

8. Serre's Conjecture

In Wiles' original lectures in Cambridge in 1993, he introduced his method with the statement that it was *orthogonal* to Serre's conjecture [150]. In some senses, this viewpoint turned out to be the opposite of prophetic, in that the ultimate resolution of Serre's conjecture used the Taylor–Wiles method as its central core. The proof of Serre's conjecture by Khare and Wintenberger [111, 112, 113, 114] introduced a new technique for lifting residual Galois representations to characteristic zero (see §8.2) which has proved very useful for subsequent modularity lifting theorems.

8.1. **Ramakrishna lifting.** Ramakrishna, in a series of papers in the late 90s [147, 146], studied the question of lifting an odd Galois representation

$$\overline{\rho} : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

to a global potentially semistable representation in characteristic zero unramified outside finitely many primes. This is a trivial consequence of Serre's conjecture[23] but is highly non-obvious without such an assumption. Ramakrishna succeeded in proving the existence of lifts by an ingenious argument involving adding auxiliary primes and modifying the local deformation problem to a setting where there all global obstructions vanished. The resulting lifts had the added property that they were valued in $\mathrm{GL}_2(W(k))$ whenever $\overline{\rho}$ was valued in $\mathrm{GL}_2(k)$. Adaptations of Ramakrishna's method had a number of important applications even under the assumption of residual modularity, including in [50] where it was used to produce characteristic zero lifts with Steinberg conditions at some auxiliary primes. There is also recent work of Fakhruddin, Khare, and Patrikis [79] which considerably extends these results in a number of directions.

8.2. **The Khare–Wintenberger method.** One disadvantage of Ramakrishna's method was that it required allowing auxiliary ramification which (assuming Serre's conjecture) should not be necessary[24]. Khare and Wintenberger found a new powerful method for

[22]The assumption on the Hodge–Tate weights is almost certainly removable using recent progress on the ideas discussed in §4 (Sasaki has announced such a result). Moreover, Pan has found a different approach to this case as well, see [138, Theorem 1.0.5] and the subsequent comments. The hypothesis that ρ is odd is more troublesome — see §9.7.

[23]Trivial only assuming the results of Tsuji [175] and Saito [151], of course.

[24]If one insists on finding a lift is valued in $\mathrm{GL}_2(W(k))$ rather than $\mathrm{GL}_2(\mathcal{O}_E)$ for some ramified extension $E/W(k)[1/p]$, then some auxiliary ramification is necessary in general, at least in fixed weight.

avoiding this. The starting point is the idea that, given an odd representation $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ for a totally real field F satisfying the Taylor–Wiles hypotheses, one could find a finite extension H/F where $\overline{\rho}$ is modular (exactly as in §5). Then, using an $R = \mathbf{T}$ theorem over H, one proves that the corresponding deformation ring R_H of $\overline{\rho}|_{G_H}$ is finite over $W(k)$. However, for formal reasons, there is a map $R_H \to R_F$ (where R is the deformation ring corresponding to the original representation $\overline{\rho}$) which is a finite morphism, and hence the ring R_F/p is Artinian. Then, by Galois cohomological arguments, one proves the ring R_F has dimension at least one, from which one deduces that R_F has $\overline{\mathbf{Q}}_p$-valued points. Even more can be extracted from this argument, however — the $\overline{\mathbf{Q}}_p$-valued point of R_F certainly comes from a $\overline{\mathbf{Q}}_p$-valued point of R_H, and hence comes from a compatible family of Galois representations over H. Using the fact that one member of the family extends to G_F, it can be argued that the entire family descends to a compatible family over F. This one can then hope to prove is modular by working at a different (possibly smaller) prime, where (hopefully) one can prove the associated residual representation is modular. In this way, one can inductively reduce Serre's conjecture [156] to the case $p = 2$ and $N(\overline{\rho}) = 1$, where Tate had previously proved in a letter to Serre [62, July 2, 1973] (also [164]) that all such absolutely irreducible representations are modular by showing that no such representations exist. The entire idea is very clean, although in practice the difficulty reduces to the step of proving modularity lifting theorems knowing either that $\overline{\rho}$ is either modular and absolutely irreducible or is reducible. Khare and Wintenberger's timing was such that the automorphy lifting technology was just good enough for the proof to work, although this required some extra effort at the prime $p = 2$ (both in their own work and in a key assist by Kisin [121]). As with Ramakrishna's method, the Khare–Wintenberger lifting method has also been systematically exploited for modularity lifting applications (for example in [11] (see §9.6) building on ideas of Gee [87]).

9. Higher dimensions

Parallel to the developments of p-adic Langlands for $n = 2$, the first steps were made to generalize the theory to higher dimensional representations. Unlike in the case of modular forms, substantially less was known about the existence of Galois representations until the 90s.

9.1. **Construction of Galois representations, part I: Clozel–Kottwitz.** The first general construction[25] of Galois representations in dimension $n > 2$ was made by Clozel [48] (see also the work of Kottwitz [125]). Clozel's theorem applies to certain automorphic forms for $\mathrm{GL}_n(\mathbf{A}_L)$ for CM fields L/L^+. The construction requires three important hypotheses on π: The first is that π is conjugate self-dual, that is $\pi^\vee \simeq \pi^c$. If π is a base change from

[25] Clozel's paper is from 1991 and thus not strictly "post–Fermat" as is the remit of this survey. However, it can be considered a natural starting point for the "modern" arithmetic theory of automorphic forms for $\mathrm{GL}(n)$ and so it seems reasonable to mention it here.

an algebraic representation of L^+ and $n = 2$ then this condition is automatic[26], but it is far from automatic when $n > 2$. The second condition is an assumption on the infinitesimal character which (in the case of modular forms) is equivalent to the condition that the weight k is ≥ 2. Finally, there is a technical condition that for some finite place x the representation π_x is square integrable. A number of improvements (particularly at the bad primes) were made by Harris–Taylor in [100, Theorem C], and later by Taylor–Yoshida and Caraiani [170, 42, 43], bringing the theory roughly in line with that of modular forms at the time of Wiles, and in particular primed for possible generalizations of the Taylor–Wiles method to higher dimensions.

9.2. **The Sato–Tate conjecture, part I.** Harris, and Taylor (as early as 1996) started the work of generalizing the Taylor–Wiles machinery to the setting of n-dimensional representations. They quickly understood that the natural generalization of these ideas in n-dimensions required the hypothesis that the Galois representations were self-dual up to a twist. This meant that one should not consider general automorphic forms on the group $\mathrm{GL}_n(\mathbf{A}_{\mathbf{Q}})$ but rather groups of symplectic or orthogonal type depending on the parity of n. If one replaced Galois representations over totally real fields by Galois representations over imaginary CM fields and then further imposed the condition that the Galois representations are conjugate self-dual, the relevant automorphic forms should then come from unitary groups. There were two benefits of working with these hypotheses. First of all, the relevant automorphic representations for unitary groups were, as with modular forms, associated to cohomology classes on Shimura varieties. In particular, under the assumption that there existed an auxiliary prime x such that π_x was square integrable, they could be seen inside the "simple" Shimura varieties of type $U(n-1,1)$ considered by Kottwitz [125]. On the other hand, the same Hecke eigenclasses (if not Galois representations) also came from a compact form of the group and thus inside the cohomology of zero-dimensional varieties[27]. The advantage of working in this setting is that the freeness of M_Q over the ring of diamond operators is immediate[28]. In the fundamental paper [50], Clozel, Harris, and Taylor succeeded in overcoming many of the technical difficulties generalizing the arguments of [180, 169] to these representations. Although the argument in spirit was very much the same, there are a number of points for GL_2 where things are much easier. One representative example of this phenomenon is understanding Taylor–Wiles primes. While the Galois side generalizes readily, the automorphic side requires many new ideas and some quite subtle arguments concerning the mod p structure of certain $\mathrm{GL}_n(\mathbf{Q}_q)$-representations of conductor 1 and conductor q. In order to prove the Sato–Tate conjecture for a modular form f, it was already observed by Langlands that it sufficed to prove the modularity of all the symmetric powers of f. However, it turns out that the weaker assumption that each

[26] At least after a twist which is always possible to achieve in practice, see [50, Lemma 4.1.4]. More generally, one can work with unitary similitude groups and consider π with $\pi^\vee \simeq \pi^c \otimes \chi$ for suitable characters χ.

[27] Inside H^0, of course

[28] In more general contexts, the freeness of M_Q is closely related to the vanishing of cohomology localized at $\mathfrak{m}$ in all but one degree.

of these symmetric powers is *potentially modular* suffices, and by some subterfuge only the even powers are required [95]. In order to prove potential modularity theorems, one needs to be able to carry out some version of the p-q switch (§5). In order to do *this*, one needs a source of motives which both generate Galois representations of the right shape (conjugate self dual and with distinct Hodge–Tate weights) and yet also come in positive dimensional families. It turned out that there already existed such motives in the literature, namely, the so-called Dwork family. However, given the strength of the automorphy lifting theorems in [50], considerable effort had to be made in studying the geometry of the Dwork family to ensure that the p-q switch would produce geometric Galois representations with the right local properties. These issues were precisely addressed in the companion paper by Harris, Shepherd-Barron and Taylor [97]. Taken together, these papers contained all the ingredients to prove the potential modularity of higher symmetric powers of modular forms (satisfying a technical square integrable condition at some auxiliary prime) with one exception. As mentioned earlier, the work of Kisin had simplified the passage from the minimal case to the non-minimal case — "all" that was required was to produce congruences between the original form and forms of higher level rather than to compute a precise congruence number as in [180]. However, even applying Kisin's approach seemed to require Ihara's Lemma, and despite several years of effort, the authors of [50] were not able to overcome this obstacle[29]. Here is Michael Harris' recollection of the process:

> In the spring of 1995, I was at Brandeis, Richard was at MIT, and I wanted to understand the brand new proof of Fermat's Last Theorem. So I asked Richard if he would help me learn by collaborating on modularity for higher-dimensional groups. The collaboration took off a year later, when Richard wrote to tell me about the Diamond–Fujiwara argument and suggested that we work out the Taylor–Wiles method for unitary groups. This developed over the next 18 months or so into the early version of what eventually became the IHES paper with Clozel. But it had no punch line. I was hoping to work out some non-trivial examples of tensor product functoriality for $\mathrm{GL}(n) \times \mathrm{GL}(m)$, where one of the two representations was congruent mod l to one induced from a CM Hecke character. This would have required some numerical verification. In the meantime we got sidetracked into proving the local Langlands conjecture [100].
>
> The manuscript on automorphy lifting went through several drafts and was circulated; you can still read it on my home page [99]. Genestier and Tilouine [88] quoted it when they proved modularity lifting for Siegel modular forms. When Clozel saw the draft he told me we should try to prove the Sato–Tate Conjecture. Although this was in line with my hope for examples of tensor product functoriality, it seemed completely out of reach, because I saw no way to prove residual modularity of symmetric powers.
>
> When I heard about the Skinner–Wiles paper I came up with a quixotic plan to prove symmetric power functoriality for Eisenstein representations, using the main conjecture of Iwasawa theory to control the growth of the deformation rings. This was in the spring of 2000, at the IHP special semester on the Langlands program, where I first met Chris Skinner.
>
> One day Chris told me that Richard had invented potential modularity. This led me to a slightly less hopeless plan to prove potential symmetric power functoriality by proving it for 2-dimensional representations congruent to potentially abelian representations, as in the potential modularity argument. I told Richard about this idea, probably the day he arrived in Paris. He asked: why apply potential modularity to the 2-dimensional representation; why not instead apply it to the the symmetric power representations directly? I then replied: that would require a variation of Hodge structures with a short list of properties: mainly, the correct $h^{p,q}$'s and large monodromy groups. We checked that potential modularity was

[29] The issue remains unresolved to this day.

sufficient for Sato–Tate. We then resolved to ask our contacts if they knew of VHS with the required properties. The whole conversation lasted about 20 minutes.

I asked a well-known algebraic geometer, who said he did not know of any such VHS. Richard asked Shepherd-Barron, who immediately told him about the Calabi-Yau hypersurfaces that had played such an important role in the mirror symmetry program. (And if my algebraic geometer hadn't wanted to be dismissive, for whatever reason, he would have realized this as well.) The $h^{p,q}$'s were fine but we didn't know about the monodromy. However, Richard was staying at the IHES, and by a happy accident so was Katz, and when Richard asked Katz about the monodromy for this family of hypersurfaces Katz told him they were called the Dwork family and gave him the page numbers in one of his books.

So within a week or two of our first conversation, we found ourselves needing only one more result to complete the proof of Sato–Tate. This was Ihara's lemma, which occupied our attention over the next five years. In the meantime, Clozel had written a manuscript on symmetric powers, based on the reducibility mod ell of symmetric powers. The argument was incomplete but he had several ideas that led to his joining the project, and he also hoped to use ergodic theory to prove Ihara's lemma. In the summer of 2003 Clozel and I joined Richard in "old" Cambridge to try to work this out. The rest you know. We finally released a proof conditional on Ihara's lemma in the fall of 2005. A few months later Richard found his local deformation argument, and the proof was complete.

9.3. **Taylor's trick: Ihara Avoidance.** Shortly after the preprints [50, 157] appeared, Taylor found a way to overcome the problem of Ihara's lemma. Inspired by Kisin's formulation of the Taylor–Wiles method (§6.2), Taylor had the idea of comparing two global deformation rings R^1 and R^2. Here (for simplicity) the local deformation problems associated to R^1 and R^2 are formally smooth at all but a single prime q. At the prime q, however, the local deformation problem associated to R^1 consists of tamely ramified representations where a generator σ of tame inertia has characteristic polynomial $(X-1)^n$, and for R^2 the characteristic polynomial has the shape $(X-\zeta_1)\ldots(X-\zeta_n)$ for some fixed distinct roots of unity $\zeta_i \equiv 1 \bmod p$. On the automorphic side, there are two patched modules H_1 and H_2, and there is an equality $H_1/p = H_2/p$. The local deformation ring R^1_q associated to R^1 at q is reducible and has multiple components in the generic fibre, although the components in characteristic zero are in bijection to the components in the special fibre. On the other hand, the local deformation ring associated to R^2 at q consists of a single component, and so using Kisin's argument one deduces that H_2 has full support. Now a commutative algebra argument using the identity $H_1/p = H_2/p$ and the structure of R^1_q implies that H_1 has sufficiently large support over R_1, enabling one to deduce the modularity of every $\overline{\mathbf{Q}}_p$-valued point of R_1[30].

9.4. **The Sato–Tate conjecture, part II.** After Taylor's trick, one was *almost* in a position to complete the proof of Sato–Tate for all classical modular forms. A few more arguments were required. One was the tensor product trick due to Harris which enabled one to pass from conjugate self-dual motives with weights in an arithmetic progression to conjugate self-dual motives with consecutive Hodge–Tate weights by a judicious twisting argument using CM characters. A second ingredient was the analysis of the ordinary

[30]Taylor's argument proves theorems of the form $R[1/p]^{\mathrm{red}} = \mathbf{T}[1/p]$ rather than $R = \mathbf{T}$. This is still perfectly sufficient for proving modularity lifting results, but not always other interesting corollaries associated to $R = \mathbf{T}$ theorems like finiteness of the corresponding adjoint Selmer groups (though see [3, 133]).

deformation ring by Geraghty [89]. One of the requirements of the p-q trick was the condition that certain moduli spaces (the Dwork family in this case) had points over various local extensions E of $\mathbf{Q}_p$, in order to construct a motive M over a number field F with $F_v = E$ for $v|p$. For the purposes of modularity lifting, one wants strong control over the local deformation ring at p, and the choice of local deformation ring is more or less forced by the geometric properties of the p-adic representations associated to M . One way to achieve this would be to work in the Fontaine–Laffaille range where the local deformation rings were smooth. But this requires both that M is smooth at p and that the ramification degree e of $E/\mathbf{Q}_p$ is one. It is not so clear, however, that the Dwork family contains suitable points (for a fixed residual representation $\overline{\rho}$) which lie in any unramified extension of $\mathbf{Q}_p$. What Geraghty showed, however, was that certain ordinary deformation rings[31] were connected over arbitrarily ramified bases. The final piece, however, was the construction of Galois representations for all conjugate self-dual regular algebraic cuspidal π without the extra condition that π_q was square integrable for some q. This story merits it own separate discussion; suffice to say that it required the combined efforts of many people and the resolution of many difficult problems, not least of which was the fundamental lemma by Laumon and Ngô [128, 136] (see also the Paris book project [49, 93], the work of Shin [158], and many more references which if I attempted to make complete would weigh down the bibliography and still contain grievous omissions).

9.5. **Big image conditions.** The original arguments in [180, 169] required a "big image" hypothesis, namely that $\overline{\rho}$ was absolutely irreducible after restriction to the Galois group of $\mathbf{Q}(\zeta_p)$. Wiles' argument also required the vanishing of certain cohomology groups associated to the adjoint representation of the image of $\overline{\rho}$. These assumptions had natural analogues in [50] (so-called "big image" hypotheses) although they were quite restrictive, and it wasn't clear that they would even apply to most residual representations coming from some irreducible compatible family. In the setting of 2-dimensional representations, the Taylor–Wiles hypothesis guarantees the existence of many primes q such that $q \equiv 1 \bmod p$ and such that $\overline{\rho}(\mathrm{Frob}_q)$ has distinct eigenvalues. This ensures, for example, that there cannot be any Steinberg deformations at q because the ratio of the eigenvalues of any Steinberg deformation must be q. In dimension n, one natural way to generalize this might be to say that $\overline{\rho}(\mathrm{Frob}_q)$ has distinct eigenvalues, although this is not always possible to achieve for many irreducible representations $\overline{\rho}$. A weaker condition is that $\overline{\rho}(\mathrm{Frob}_q)$ has an eigenvalue α with multiplicity one. For such q, there will be no deformations which are unipotent on inertia at q for which the generalized α eigenspace is not associated to a 1-dimensional block. The translation of this into an automorphic condition on U_q-eigenvalues is precisely what is done in [50] (there are additional technical conditions on Frob_q with respect to the adjoint representation $\mathrm{ad}(\overline{\rho})$ which we omit here). In [172], however, Thorne finds a way to allow $\overline{\rho}(\mathrm{Frob}_q)$ to have an eigenvalue α with higher multiplicity, and yet still

[31]In Geraghty's setting, the residual representations $\overline{\rho}$ were locally trivial. Hence the definition of "ordinary" was not something that could be defined on the level of Artinian rings, and the construction (as with Kisin's construction of local deformation rings associated to certain types) is therefore indirect.

cut out (integrally) the space of automorphic forms whose Galois representations decompose at q as an unramified representation plus a one dimensional representation which is tamely ramified of p-power order. This technical improvement is very important because (as proved in the appendix by Guralnick, Herzig, Taylor, and Thorne [172]) it imposes no restrictions on $\overline{\rho}$ when $p \geq 2n+1$ beyond the condition that $\overline{\rho}$ is absolutely irreducible after restriction to $G_{\mathbf{Q}(\zeta_p)}$. This improvement is very useful for applications.

9.6. **Potentially diagonalizable representations.** After the proof of Sato–Tate for modular forms, Barnet-Lamb, Gee, and Geraghty turned their attention to proving the analogous theorem for Hilbert modular forms of regular weight. The methods developed so far were well-suited both to representations ρ which were either ordinary or when ρ was not ordinary but still Fontaine–Laffaille. (The latter implies that $\overline{\rho}|_{G_{\mathbf{Q}_p}}$ is absolutely irreducible of some particular shape.) For a modular form over $\mathbf{Q}$, one easily sees that ρ takes one of one of these forms for any sufficiently large p. For Hilbert modular forms, one certainly expects that the ordinary hypothesis should hold for all $v|p$ and infinitely many p, but this remains open. The difficulty arises when, for some prime p (that splits completely, say) the p-adic representation is ordinary at some $v|p$ but non-ordinary other $v|p$. The reason that this causes issues is that, when applying the Moret-Bailly argument in the p-q switch, one wants to avoid any ramification at p for the non-ordinary case, and yet have large ramification at the ordinary case to make $\overline{\rho}$ locally trivial, and these desires are not compatible. The resolution in [10] involved a clever refinement of the Harris tensor product trick. These ideas were further refined in [11] and led to the concept of a *potentially diagonalizable* representation $\rho : G_E \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p)$ for some finite extension of $E/\mathbf{Q}_p$. Recall from §6.2 that, in the modified form of the Taylor–Wiles method, proving modularity of some lift of $\overline{\rho}$ often comes down to showing the existence of a modular lift lying on a smooth point of the corresponding component of the generic fibre of R^{loc}. In light of Taylor's Ihara avoidance trick (§9.3), the difficulty in this problem is mostly at the prime p, and in particular the fact that one knows very little about the components of general Kisin potentially crystalline deformation rings. A potentially diagonalizable representation is one for which, after some finite (necessarily solvable!) extension E'/E, the representation $\rho|_{G'_E}$ is crystalline and lies on the same generic irreducible component as a diagonal representation. This notion has a number of felicitous properties. First, it includes Fontaine–Laffaille representations and ordinary potentially crystalline representations. Second, it is clearly invariant under base change. Third, it is compatible with the tensor product trick of Harris. These features make it supremely well-adapted to the current forms of the Taylor–Wiles method. By combining this notion with methods of [10, 12], as well as extensive use of Khare–Wintenberger lifting (§8.2), Barnet–Lamb, Gee, Geraghty, and Taylor in [11] proved the potential automorphy of *all* conjugate self-dual irreducible[32] odd[33] compatible systems of Galois representations over a totally real field.

9.7. **Even Galois Representations.** The Fontaine–Mazur conjecture for geometric Galois representations $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$ predicts that, up to twist, either ρ is modular *or* ρ is even with finite image. The methods of [180, 169] required the assumption that $\overline{\rho}$ was modular and so *a priori* the assumption that ρ was odd (at least when $p > 2$). Nothing at all was known about the even case before the papers [34, 35] in which a very simple trick made the problem accessible to modularity lifting machinery under the assumption that the Hodge–Tate weights are distinct. The punch line is that, for any CM field F/F^+, the restriction $\mathrm{Sym}^2(\rho) : G_F \to \mathrm{GL}_3(\overline{\mathbf{Q}}_p)$ is conjugate self-dual and no longer sees the "evenness" of ρ[34]. Hence one can hope to prove it is potentially modular for some CM extension L/L^+, and then by cyclic base change [5] potentially modular for the totally real field L^+. But Galois representations coming from regular algebraic automorphic forms for totally real fields will not be even[35], and thus one obtains a contradiction. These ideas are already enough to deduce the main result of [34] directly from [11], although in contrast [35] uses (indirectly) the full strength of the p-adic local Langlands correspondence via theorems of Kisin [120]. The papers [34, 35] still fall short of completely resolving the Fontaine–Mazur in this case even for $p > 7$, since there remain big image hypotheses on $\overline{\rho}$. On the other hand, this trick has nothing to say about the case when the Hodge–Tate weights are equal (see §12).

9.8. **Modularity of higher symmetric powers.** Another parallel development in higher dimensions was the extension of Skinner–Wiles (§3) to higher dimensions. Many of the arguments of Skinner–Wiles relied heavily on the fact that any proper submodule of a 2-dimensional representation must have dimension 1, and one-dimensional representations are very well understood by class field theory. Nonetheless, in [173], Thorne proved a residually reducible modularity theorem for higher dimensional representations. In order to overcome the difficulty of controlling reducible deformations, he imposed a Steinberg condition at some auxiliary prime. Although this is a definite restriction, it does apply (for example) to the Galois representation coming from the symmetric power of a modular form which also satisfies this condition. In a sequence of papers [51, 52, 53], Clozel and Thorne applied this modularity lifting theorem to prove new cases of symmetric power functoriality (see also the paper of Dieulefait [74]). A key difficulty here is again the absence of Ihara's

[32]One variant proved shortly thereafter by Patrikis–Taylor [139] replaced the irreducibility condition by a purity condition (which is automatically satisfied by representations coming from pure motives).

[33]Although there is no longer a non-trivial complex conjugation in the Galois group of a CM field, there is still an oddness condition related to the conjugate self-duality of the representation and the fact that there are two ways for an irreducible representation to be self-dual (orthogonal and symplectic).

[34]The representation ρ itself restricted to G_F will not be odd in the required sense — one exploits the fact here that 3 is odd whereas symplectic representations are always even dimensional.

[35]I managed to twist Taylor's arm into writing the paper [168] which proved this for odd n, which sufficed for my purposes where n was either 3 or 9. This is now also known for general n, see Caraiani–Le Hung [44].

lemma in order to find automorphic forms with the correct local properties. Very recently (using a number of new ideas), Newton and Thorne [134, 135] were able to (spectacularly!) complete this program and prove the full modularity of all symmetric powers of all modular forms.

10. Beyond self-duality and Shimura varieties

All the results discussed so far — with the exception of those discussed in §4 —- apply only to Galois representations which are both regular and satisfy some form of self-duality. Moreover, they all correspond to automorphic forms which can be detected by the (étale) cohomology of Shimura varieties. Once one goes beyond these representations, many of the established methods begin to break immediately[36].

An instructive case to consider is the case of 2-dimensional geometric Galois representations of an imaginary quadratic field F with distinct Hodge–Tate weights. The corresponding automorphic forms for $\mathrm{GL}_2(\mathbf{A}_F)$ contribute to the cohomology of locally symmetric spaces X which are arithmetic hyperbolic 3-manifolds[37]. These spaces are certainly not algebraic varieties and their cohomology is hard to access via algebraic methods. One of the first new questions to arise in this context is the relationship between torsion classes and Galois representations. Some speculations about this matter were made by Elstrodt, Grunewald, and Mennicke at least as far back as 1981 [75], but the most influential conjecture was due to Ash [6], who conjectured that eigenclasses in the cohomology of congruence subgroups of $\mathrm{GL}_n(\mathbf{Z})$ over $\overline{\mathbf{F}}_p$ (which need not lift to characteristic zero) should give rise to n-dimensional Galois representations over finite fields. Later, conjectures were made [8, 7] in the converse conjecture in the spirit of Serre [156] linking Galois representations to classes in cohomology modulo p. Certainly around 2004, however, it was not at all clear what exactly one should expect the landscape to be[38], and so it was around this time I decided to start thinking about this question[39] in earnest. I became convinced very soon (for aesthetic reasons if not anything else) that if one modified $\mathbf{T}$ to be the ring of endomorphisms acting on integral cohomology (so that it would see not only

[36] I should warn the reader that this section and the next (even more than the rest of this paper) is filtered through the lens of my own personal research journey — *caveat lector*!

[37] Already by 1970, Serre (following ideas of Langlands) was trying to link Mennicke's computation that $\mathrm{GL}_2(\mathbf{Z}[\sqrt{-109}])^{\mathrm{ab}}$ is infinite to the possible existence of elliptic curves over $\mathbf{Q}(\sqrt{-109})$ with good reduction everywhere [61, Jan 14, 1970].

[38] I recall conversations with a number of experts at the 2004 Durham conference, where nobody seemed quite sure even what the dimension of the ordinary deformation ring R of a 3-dimensional representation $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_3(\overline{\mathbf{F}}_p)$ should be. Ash, Pollack, and Stevens had computed numerical examples where a regular algebraic ordinary cuspidal form for $\mathrm{GL}_3(\mathbf{A}_{\mathbf{Q}})$ not twist-equivalent to a symmetric square did not appear to admit classical deformations. (I learnt about this example from Stevens at a talk at Banff in December 2003.) This would be easily explained if R had (relative) dimension 0 over $\mathbf{Z}_p$ but be more mysterious otherwise.

[39] One great benefit to me at the time of thinking about Galois representations over imaginary quadratic fields was that it didn't require me to understand the geometry of Shimura varieties which I have always found too complicated to understand. The irony of course is that the results of [1, 17] ultimately rely on extremely intricate properties of Shimura varieties.

the relevant automorphic forms but also the torsion classes) then there should still be an isomorphism $R = \mathbf{T}$. Moreover, this equality would not only be a form of reciprocity which moved beyond the conjecture linking motives to automorphic forms, but it suggested that the integral cohomology of arithmetic groups (including the torsion classes) were themselves the fundamental object of interest. Various developments served only to confirm this point of view. In my paper with Mazur [40], we gave some theoretical evidence for why ordinary families of Galois representations of imaginary quadratic fields might on the one hand be positive dimensional and explained completely by torsion classes and yet not contain any classical automorphic points at all. During the process of writing [36], Dunfield (numerically) compared the torsion classes in the cohomology of inner forms of GL_2 and the data was in perfect agreement with a conjectural Jacquet–Langlands correspondence for torsion (later taken up in joint work with Venkatesh [41]). Emerton and I had the idea of working with completed cohomology groups both to construct Galois representations and even possibly to approach questions of modularity. The first idea was to exploit the well-known relationship between the cohomology of these manifolds and the cohomology of the boundary of certain Shimura varieties. We realized that if we could control the co-dimension of the completed cohomology groups over the non-commutative Iwasawa algebra, the Hecke eigenclasses would be forced to be seen by eigenclasses coming from the middle degree of these Shimura varieties where one had access to Galois representations[40]. On the automorphy lifting side, we had even vaguer ideas [37, §1.8][41] on how to proceed. A different (and similarly unsuccessful) approach[42] was to work with ordinary deformations over a partial weight space for a split prime $p = vw$ in an imaginary quadratic field F. That is, deformations of $\overline{\rho}$ which had an unramified quotient at v and w but with varying weight at v and fixed weight at w. Here the yoga of Galois deformations suggested that R should be finite flat over $W(k)$ in this case (and even a complete intersection). Moreover, one had access to $\mathbf{T}$ using an overlooked[43] result of Hida [103], and in particular one could deduce that $\mathbf{T}$ has dimension at least one. If one could show that $\mathbf{T}$ was flat over $W(k)$, then one could plausibly apply (assuming the existence of Galois representations) the original argument of [180, 169]. The flatness of $\mathbf{T}$, however, remains an open problem[44].

10.1. **The Taylor–Wiles method when $l_0 > 0$, part I: Calegari–Geraghty.** Shortly before (and then during) the special year on Galois representations at the IAS in 2010-2011, I started to work with Geraghty in earnest on the problem of proving $R = \mathbf{T}$ in the case of

[40]Unfortunately, these conjectures [37, Conj 1.5] remain all open in more or less all cases except for Scholze's results in the case of certain Shimura varieties [155, Cor 4.2.3].

[41]Pan's remarkable paper [137] turned some of these pipe dreams into reality.

[42]This is taken from my 2006 NSF proposal, and I believe influenced by my conversations with Taylor at Harvard around that time.

[43]One should never overlook results of Hida. I only learnt about this paper when Hida pointed it out to me (with a characteristic smile on his face) after my talk in Montreal in 2005. I was pleased at least that the idea that these families were genuinely non-classical was not anticipated either in [103] or in §4 of Taylor's thesis [171].

[44]One might even argue that there is no compelling argument to believe it is true — the problem is analogous to the vanishing of the μ-invariant in Iwasawa theoretic settings.

imaginary quadratic fields, assuming the existence of a surjection $R \to \mathbf{T}$. A computation in Galois cohomology shows that the expected "virtual" dimension of R over $W(k)$ should be -1, and hence the patched module M_∞ should have codimension 1 over the ring of diamond operators S_∞. We realized this was a consequence of the fact that, after localizing the cohomology at a non-Eisenstein maximal ideal, the cohomology should be non-zero in exactly two degrees. More precisely, patching the presentations of these S_N-modules would result in a balanced presentation of M_∞ as an S_∞-module with the same (finite) number of generators as relations. We then realized that the same principle held more generally for n-dimensional representations over any number field. In characteristic zero, the localized cohomology groups were non-zero exactly in a range $[q_0, q_0 + l_0]$ (with q_0 and l_0 as defined in [26]) where $-l_0$ coincided with the expected virtual dimension of R over $W(k)$ coming from Galois cohomology. We could thus show — assuming the localized torsion cohomology also vanished in this range — that by patching *complexes* P_Q (rather than modules M_Q), one arrives at a complex P_∞ of free S_∞ modules in degrees $[q_0, q_0 + l_0]$. Because the ring R_∞ of dimension $\dim R_\infty = (\dim S_\infty) - l_0$ acts by patching on $H^*(P_\infty)$, a simple commutative algebra lemma then shows that $M_\infty = H^*(P_\infty)$ has codimension l_0 over S_∞ and must be concentrated in the final degree. In particular, the Taylor–Wiles method (as modified by Diamond) could be happily adapted to this general setting[45]. Moreover, the arguments were compatible with all the other improvements, including Taylor's Ihara avoidance argument amongst other things[46]. We also realized that the same idea applied to Galois representations coming from the coherent cohomology of Shimura varieties even when the corresponding automorphic forms were not discrete series. While our general formulation involved a number of conjectures we considered hopeless, the coherent case had at least one setting in which many more results were available, namely the case of modular forms of weight one, where the required vanishing conjecture was obvious, and where we were able to establish the existence of the required map $R \to \mathbf{T}$ with all the required local properties by direct arguments. Although the state of knowledge concerning Galois representations increased tremendously between the original conception of [38] and its final publication, by early 2016 it still seemed out of reach to make any of the results in [38] unconditional.

10.2. **Construction of Galois representations, part II.** Before one can hope to prove $R = \mathbf{T}$ theorems, one needs to be able to associate Galois representations to the corresponding automorphic forms. There are two contexts in which one might hope to make progress. The first is in situations where the automorphic forms contribute to the Betti cohomology of some locally symmetric space — for example, tempered algebraic cuspidal automorphic

[45]David Hansen came up with a number of these ideas independently [94].

[46]These methods only prove $R[1/p]^{\mathrm{red}} = \mathbf{T}[1/p]^{\mathrm{red}}$, of course. In situations where $\mathbf{T} \otimes \mathbf{Q} = 0$, the methods of [38] in the minimal case prove not only that $R = \mathbf{T}$ but also that (both) rings are complete intersections. Moreover, one also has access to level raising (on the level of complexes) and Ihara's lemma [41, §4], and I tried for some time (unsuccessfully) to adapt the original minimal $\Rightarrow$ non-minimal arguments of [180] to this setting. There certainly seems to be some rich ideas in commutative algebra in these situations to explore, see for example recent work of Tilouine–Urban [174].

representations for $\mathrm{GL}_n(\mathbf{A}_F)$ and any F. The second is in situations where the automorphic forms contribute to the coherent cohomology of some Shimura variety. Here the first and easiest case corresponds to weight one modular forms, where the Galois representations were first constructed by Deligne and Serre [69].

In work of Harris–Soudry–Taylor [98, 165], Galois representations were constructed for regular algebraic forms for $\mathrm{GL}_2(\mathbf{A}_F)$ for an imaginary quadratic field F and satisfying a further restriction on the central character. Harris, Soudry, and Taylor exploited (more or less) the fact that the automorphic induction of such forms are self-dual (although not regular) and still contribute to coherent cohomology, so one can construct Galois representations using a congruence argument as in the paper of Deligne and Serre [69]. On the other hand, this does not prove the expected local properties of the Galois representation at $v|p$.

It was well-known for many years that the Hecke eigenclasses associated to regular algebraic cuspidal automorphic forms for $\mathrm{GL}_n(\mathbf{A}_F)$ for a CM field F could be realized as eigenclasses coming from the boundary of certain unitary Shimura varieties of type $U(n, n)$. It was, however, also well-known that the corresponding étale cohomology classes did not realize the desired Galois representations[47]. Remarkably, this problem was completely and unexpectedly resolved in 2011 in [96] by Harris–Lan–Taylor–Thorne. Richard Taylor writes:

> For [96] I knew that the Hecke eigenvalues we were interested in contributed to Betti cohomology of $U(n, n)$. The problem was to show that they contributed to overconvergent p-adic cusp forms. I was convinced on the basis of Coleman's paper "classical and overconvergent modular forms" [55] that this must be so. I can't now reconstruct exactly why Coleman's paper convinced me of this, and it is possible, even probable, that my reasoning didn't really make any sense. However it was definitely this that kept me working at the problem, when we weren't really getting anywhere.

Amazingly, this breakthrough immediately inspired the next development:

10.3. **Construction of Galois representations, part III: Scholze.** In [155], Scholze succeeded in constructing Galois representations associated to torsion classes in the setting of $\mathrm{GL}_n(\mathbf{A}_F)$ for a CM field F. Scholze had the idea after seeing some lectures on [96]:

> During a HIM trimester at Bonn, Harris and Lan gave some talks about their construction of Galois representations. At the time, I had some ideas in my head that I didn't have any use for: That Shimura varieties became perfectoid at infinite level, and that there is a Hodge–Tate period map defined on them. The only consequence I could draw from this were certain vanishing results for completed cohomology as conjectured in your work with Emerton; so at least I knew that the methods were able to say something nontrivial about torsion classes in the cohomology. After hearing Harris' and Lan's talks, I was trying to see whether these ideas could help in extending their results to torsion classes. After a little bit of trying, I found the fake-Hasse invariants, and then it was clear how the argument would go.

Even after this breakthrough, Scholze's construction still fell short of the conjectures in [38] in two ways. The first was that the Galois representation (ignoring here issues of pseudo-representations) was valued not in $\mathbf{T}$ but in $\mathbf{T}/I$ for some ideal I of fixed nilpotence. This is not a crucial obstruction to the methods of [38]. The second issue, however, was that

[47] For a more basic example of what can go wrong, note that the Hecke eigenvalues of T_l on $H^0(X, \mathbf{Q}_p)$ of a modular curve are $1 + l$, which corresponds to the Galois representation $\mathbf{Q}_p \oplus \mathbf{Q}_p(1)$. However, only the piece $\mathbf{Q}_p$ occurs inside H^0.

the Galois representations were constructed (in the end) via p-adic congruences, and thus one did not have control over their local properties at p which *are* crucial for modularity applications.

10.4. **The Taylor–Wiles method when $l_0 > 0$, part II: DAG.** Although not directly related to new $R = \mathbf{T}$ theorems, one new recent idea in the subject has been the work of Galatius–Venkatesh [86] on derived deformation rings in the context of Venkatesh's conjectures over $\mathbf{Z}$. This work (in part) reinterprets the arguments of [38] in terms of a derived Hecke action. The authors define a derived version $\mathcal{R}$ of R with $\pi_0(\mathcal{R}) = R$. Under similar hypotheses to [38], the higher homotopy groups of $\mathcal{R}$ are shown to exist precisely in degrees 0 to l_0. One viewpoint of the minimal case of [38] is that one constructs a (highly non-canonical) formally smooth ring R_∞ of dimension $n - l_0$ with an action of a formally smooth ring S_∞ of dimension n such that the minimal deformation ring R is $R_\infty \otimes_{S_\infty} S_\infty/\mathfrak{a}$ for the augmentation ideal $\mathfrak{a}$. Moreover, the ring R is identified both with the action of $\mathbf{T}$ on the entire cohomology and simultaneously on the cohomology in degree $q_0 + l_0$. On the other hand, when $l_0 > 1$, the intersection of R_∞ and $S_\infty/\mathfrak{a}$ over S_∞ is never transverse[48], and homotopy groups of the derived intersection recover the cohomology in all degrees (under the running assumptio2n one also knows that the patched cohomology is free). On the other hand, there is a more canonical way to define $\mathcal{R}$, namely to take the unrestricted global deformation ring R^{glob} (which has no derived structure) and intersect it with a suitable local crystalline deformation ring as algebras over the unrestricted local deformation ring. The expected dimension of this intersection is also $-l_0$ over $W(k)$, although this is not so clear from this construction. Hence [86] can be viewed as giving an intrinsic definition of $\mathcal{R}$ independent of any choices of Taylor–Wiles primes and showing that its homotopy groups are related (as with $R_\infty \otimes^{\mathbf{L}}_{S_\infty} S_\infty/\mathfrak{a}$) to the cohomology[49]. These ideas have hinted at a closer connection between the Langlands program in the arithmetic case and the function field case than was previously anticipated[50], see for example work of Zhu [181].

11. Recent Progress

11.1. **Avoiding conjectures involving torsion I: the 10-author paper.** As mentioned in §10.3, even after the results of [155] there remained a significant gap to make the results of [38] unconditional, namely, the conjecture that these Galois representations had the right local properties at p and a second conjecture predicting the vanishing of

[48] When $l_0 = 1$, the intersection *can* be transverse when R is a finite ring. In this setting, the relevant cohomology is also non-zero and finite in exactly one degree. On the other hand, as soon as $\mathrm{Hom}(R, \overline{\mathbf{Q}}_p)$ is non-zero (for example, when there exists an associated motive) and $l_0 > 0$, the intersection will always be non-transverse.

[49] There are some subtleties as to what the precise statement should be in the presence of global congruences, but already this author gets confused at the best of times between homology and cohomology, so I will not try to unentangle these issues here.

[50] Not anticipated by many people, at least; Michael Harris has been proselytizing the existence of a connection for quite some time.

(integral) cohomology localized at a non-maximal ideal $\mathfrak{m}$ outside a certain precise range (corresponding to known results in characteristic zero). It should be noted that the second conjecture was still open (in all but the easiest cases) in the simpler setting of Shimura varieties. The first hints that one could possibly make progress on this second conjecture (at least for Shimura varieties) was given in an informal talk by Scholze in Bellairs[51] in 2014. This very quickly led to a long term collaboration between Scholze and Caraiani [45, 46], which Caraiani describes as follows:

> At the Barbados conference in May 2014, Peter gave a lecture on how one might compute the cohomology of compact unitary Shimura varieties with torsion coefficients. The key was to have some control for $R\pi_{HT*}\mathbb{F}_\ell$ restricted to any given Newton stratum. He was expressing this in terms of a conjecture that had grown out of his work on local Langlands using the Langlands–Kottwitz method. After his talk, I went to ask him some questions about this conjecture and it sounded like there were some things that still needed to be made precise. He asked if I wanted to help him make his strategy work. After some hesitation (because I didn't think I knew enough or was strong enough to work with him), I accepted. Later that evening, I suggested switching from the Langlands–Kottwitz approach to understanding $R\pi_{HT*}\mathbb{F}_\ell$ to an approach more in the style of Harris–Taylor. This relies on the beautiful Mantovan product formula that describes Newton strata in terms of Rapoport–Zink spaces and Igusa varieties. Maybe something like this could help illuminate the geometry of the Hodge–Tate period morphism? Peter immediately saw that this should work and we made plans for me to visit Bonn that summer to continue the collaboration.
>
> As Peter and I were finishing writing up the compact case, it became clear to us that the vanishing theorem would give a way to construct Galois representations associated to generic mod p classes that preserves the desired information at p. Peter started thinking about the non-compact case and how that might apply to the local-global compatibility needed for Calegari–Geraghty. I remember discussing this with him at the Clay Research conference in Oxford in September 2015. By spring 2016, Richard started floating the idea of a working group on Calegari–Geraghty and found out that Peter and I had an approach to local-global compatibility. Around June 2016, Richard suggested to me to organize the working group with him. Peter was very excited about the idea, but wasn't sure he would be able to attend for family reasons. In the end, we found a date in late October 2016 that worked for everyone.

The working group met under the auspices of the first "emerging topics" workshop[52] at the IAS to determine the extent to which the expected consequences could be applied to modularity lifting: A clear stumbling point was the vanishing of integral cohomology after localization outside the range of degrees $[q_0, q_0 + l_0]$. On the other hand, Khare and Thorne had already observed in [115] by a localization argument that this could sometimes be avoided in certain minimal cases. It was this argument we were able to modify for the general case, thus avoiding the need to prove the (still open) vanishing conjectures for torsion classes[53]. The result of the workshop was a success beyond what we could have

---

[51] I was invited to give the lecture series in Bellairs after Matthew Emerton didn't respond to his emails. Through some combination of the appeal of my own work and the fact that the lectures were given on a beach in Barbados, I managed to persuade Patrick Allen, George Boxer, Ana Caraiani, Toby Gee, Vincent Pilloni, Peter Scholze, and Jack Thorne to come, all of whom are now my co-authors, and all of whom (if they weren't already at the time) are now more of an expert in this subject than I am. The thought that I managed to teach any of them something about the subject is pleasing indeed.

[52] Although later described as a "secret" workshop, it was an "invitation-only working group."

[53] I regard my main contribution to [1] as explaining how the arguments in [38] using Taylor's Ihara avoidance (§9.3) were incompatible with any characteristic zero localization argument in the absence of (unknown) integral vanishing results in cohomology. The objection (even in the case $l_0 = 0$) was that it was

reasonably anticipated — we ended up with more or less[54] the outline of a plan to prove all the main modularity lifting theorems which finally appeared in [1], namely the Ramanujan conjecture for regular algebraic automorphic forms for $\mathrm{GL}_2(\mathbf{A}_F)$ of weight zero for any CM field F, and potential modularity (and the Sato–Tate conjecture) for elliptic curves over CM fields.

There have already been a number of advancements beyond [1] including in particular by Allen, Khare, and Thorne [4] proving the modularity of many elliptic curves over CM fields and a potential automorphy theorem for ordinary representations by Qian [144]. It does not seem completely implausible that results of the strength of [11] for n-dimensional regular Galois representations of $G_{\mathbf{Q}}$ are within reach.

11.2. **Avoiding conjectures involving torsion II: abelian surfaces.** A second example that Geraghty and I had considered during the 2010–2011 IAS special year was the case of abelian surfaces, corresponding to low (irregular) weight Siegel modular forms of genus $g = 2$. It was clear that a key difficulty was proving the vanishing of $H^2(X, \omega^2)_{\mathfrak{m}}$ where X was a (compactified) Siegel 3-fold with good reduction at p, where $\mathfrak{m}$ is maximal ideal of the Hecke algebra corresponding to an absolutely irreducible representation, and where $\omega|_Y = \det \pi_* \Omega^1_{\mathcal{A}/Y}$ on the open moduli space $Y \subset X$ admitting a corresponding universal abelian surface $\mathcal{A}/Y$. In other irregular weights (corresponding to motives with Hodge–Tate weights $[0, 0, k-1, k-1]$ for $k \geq 4$) the vanishing of the corresponding cohomology groups was known by Lan and Suh [126]. The vanishing of $H^2(X, \omega^2)_{\mathfrak{m}}$ was more subtle, however, because the corresponding group does *not* vanish in general before localization in contrast to the previous cases. In [39], we proved a minimal modularity theorem for these higher weight representations and a minimal modularity theorem in the abelian case contingent on the vanishing conjecture above which we did not manage to resolve (and which remains unresolved). I finished and then submitted the paper after I had moved to Chicago and Geraghty had moved to Facebook in 2015. I then started working with Boxer and Gee[55] on this vanishing question under certain supplementary local hypotheses. (By this point, Galois representations associated to torsion classes in coherent cohomology had been constructed by Boxer [16] and Goldring–Koskivirta [90].) But then

easy to construct complexes P^1 and P^2 of free S_∞ modules so that the support of $H^*(P^1/p)$ and $H^*(P^2/p)$ coincided (as they must) but that (for example) $H^*(P^1)[1/p]$ was zero even though $H^*(P^2)[1/p]$ was not. The objection to this objection, however, which was resolved during the workshop (and which to be clear I played no part in resolving!) is to not merely to compare the support of the complexes P^i/p but to consider the entire complex in the derived category. In particular, even (say) for a finite $\mathbf{Z}_p$-module M, the module $M[1/p]$ is non-zero exactly when $M \otimes^{\mathbf{L}} \mathbf{F}_p$ has non-zero Euler characteristic.

[54]It is worth emphasizing that an incredible amount of work was required to turn these ideas into reality, and that this intellectual effort was by and large carried out by the younger members of the collaboration.

[55]George Boxer had also arrived at Chicago in 2015, and was collaborating with Gee on companion form results for Siegel modular forms, with the hope (in part) of deducing the modularity of abelian surfaces from Serre's conjecture for GSp_4 in a manner analogous to the deduction by of the Artin conjecture from Serre's conjecture for GL_2 in [110, 113]. They usually worked together at Plein Air cafe. Since I had thought about similar questions with Geraghty and frequently went to Plein Air for 6oz cappuccinos, it was not entirely surprising for us to start working together.

in November of 2016 (one week after the IAS workshop!), Pilloni's paper on higher Hida theory [142] was first posted. It was apparent to us that Pilloni's ideas would be extremely useful, and the four of us began a collaboration almost immediately. Just as in [1], we were ultimately able to avoid proving any vanishing conjectures. However, unlike [1], the way around this problem was not purely by commutative algebra, but instead by working with ideas from [142]. Namely, instead of working with the cohomology of the full Siegel modular variety X, one could work with the coherent cohomology of a certain open variety of X with cohomological dimension one whose (infinite dimensional) cohomology could still be tamed using the methods of higher Hida theory [142] in a way analogous to how Hida theory controls the (infinite dimensional) cohomology of the affine variety (with cohomological dimension zero) corresponding to the ordinary locus. Generalizing this to a totally real field, one could then *combine* these ideas with the Taylor–Wiles method as modified in [38] to prove the potential modularity of abelian surfaces over totally real fields [17]. This coincidentally gives a second proof of the potential modularity of elliptic curves over CM fields proven in [1]. (The papers [1] and [17] both were conceived of and completed within a week or so of each other.)

12. The depths of our ignorance

Despite what can reasonably be considered significant progress in proving many cases of modularity since 1993, it remains the case that many problems appear just as hopeless as they did then[56]. Perhaps most embarrassing is the case of even Galois representations $G_{\mathbf{Q}} \rightarrow \mathrm{GL}_2(\mathbf{C})$ with non-solvable image (equivalently, projective image A_5). For example, we cannot establish the Artin conjecture for a *single* Galois representation whose image is the binary icosahedral group $\mathrm{SL}_2(\mathbf{F}_5)$ of order 120. The key problem is that the automorphic forms (Maass forms with eigenvalue $\lambda = 1/4$ in this case) are very hard to access — given an even (projective) A_5 Galois representation, we don't even know how to prove that there exists a corresponding Maass form with the right Laplacian eigenvalue, let alone one whose Hecke eigenvalues correspond to the Galois representation[57]. In many ways, we have made no real progress on this question. The case of curves of genus $g > 2$ whose Jacobians have no extra endomorphisms seems equally hopeless. One can only take solace in the fact that the Shimura–Taniyama conjecture seemed equally out of reach before Wiles' announcement in Cambridge in 1993.

[56]Or perhaps harder, since there has been almost 30 years without any progress whatsoever.

[57]Motives can be divided according to a tetrachotomy. The first form are the Tate (and potentially Tate) motives, whose automorphy was known to Riemann and Hecke. The second form are the motives (conjecturally) associated to automorphic representations which are discrete series at infinity and thus amenable to the Taylor–Wiles method. The third form are the motives (conjecturally) associated to automorphic representations which are at least seen by some flavour of cohomology, either by the Betti cohomology of locally symmetric spaces or the coherent cohomology of Shimura varieties (possibly in degrees greater than zero) which are amenable in principle to the modified Taylor–Wiles method. The fourth form consist of the rest, which (besides a few that can be accessed by cyclic base change) are a complete mystery.

References

[1] P. Allen, F. Calegari, A. Caraiani, T. Gee, D. Helm, B. Le Hung, J. Newton, P. Scholze, R. Taylor, and J. Thorne, Potential automorphy over CM fields, 2018, preprint, `arXiv:1812.09999`.

[2] P. B. Allen, Modularity of nearly ordinary 2-adic residually dihedral Galois representations. *Compos. Math.* **150** (2014), no. 8, 1235–1346.

[3] P. B. Allen, Deformations of polarized automorphic Galois representations and adjoint Selmer groups. *Duke Math. J.* **165** (2016), no. 13, 2407–2460.

[4] P. B. Allen, C. Khare, and J. A. Thorne, Modularity of $\mathrm{GL}_2(\mathbb{F}_p)$-representations over CM fields. 2019, `arXiv:1910.12986`.

[5] J. Arthur and L. Clozel, *Simple algebras, base change, and the advanced theory of the trace formula.* Annals of Mathematics Studies 120, Princeton University Press, Princeton, NJ, 1989.

[6] A. Ash, Galois representations attached to mod p cohomology of $\mathrm{GL}(n, \mathbf{Z})$. *Duke Math. J.* **65** (1992), no. 2, 235–255.

[7] A. Ash, D. Doud, and D. Pollack, Galois representations with conjectural connections to arithmetic cohomology. *Duke Math. J.* **112** (2002), no. 3, 521–579.

[8] A. Ash and W. Sinnott, An analogue of Serre's conjecture for Galois representations and Hecke eigenclasses in the mod p cohomology of $\mathrm{GL}(n, \mathbf{Z})$. *Duke Math. J.* **105** (2000), no. 1, 1–24.

[9] M. Auslander and D. A. Buchsbaum, Homological dimension in local rings. *Trans. Amer. Math. Soc.* **85** (1957), 390–405.

[10] T. Barnet-Lamb, T. Gee, and D. Geraghty, The Sato-Tate conjecture for Hilbert modular forms. *J. Amer. Math. Soc.* **24** (2011), no. 2, 411–469.

[11] T. Barnet-Lamb, T. Gee, D. Geraghty, and R. Taylor, Potential automorphy and change of weight. *Ann. of Math. (2)* **179** (2014), no. 2, 501–609.

[12] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, A family of Calabi-Yau varieties and potential automorphy II. *Publ. Res. Inst. Math. Sci.* **47** (2011), no. 1, 29–98.

[13] L. Berger and C. Breuil, Towards a p-adic Langlands programme, 2004, unpublished, `http://perso.ens-lyon.fr/laurent.berger/autrestextes/hangzhou.pdf`.

[14] L. Berger and C. Breuil, Sur quelques représentations potentiellement cristallines de $\mathrm{GL}_2(\mathbf{Q}_p)$. *Astérisque* (2010), no. 330, 155–211.

[15] G. Böckle, On the density of modular points in universal deformation spaces. *Amer. J. Math.* **123** (2001), no. 5, 985–1007.

[16] G. Boxer, Torsion in the coherent cohomology of Shimura varieties and Galois representations. 2015, `arXiv:1507.05922`.

[17] G. Boxer, F. Calegari, T. Gee, and V. Pilloni, Abelian surfaces over totally real fields are potentially modular. 2018, `arXiv:1812.09269`.

[18] C. Breuil, Schémas en groupe et corps des normes, 1998, URL `https://www.imo.universite-paris-saclay.fr/~breuil/PUBLICATIONS/groupesnormes.pdf`, unpublished.

[19] C. Breuil, Groupes p-divisibles, groupes finis et modules filtrés. *Ann. of Math. (2)* **152** (2000), no. 2, 489–549.

[20] C. Breuil, Sur quelques représentations modulaires et p-adiques de $\mathrm{GL}_2(\mathbf{Q}_p)$. I. *Compositio Math.* **138** (2003), no. 2, 165–188.

[21] C. Breuil, Sur quelques représentations modulaires et p-adiques de $\mathrm{GL}_2(\mathbf{Q}_p)$. II. *J. Inst. Math. Jussieu* **2** (2003), no. 1, 23–58.

[22] C. Breuil, Introduction générale. *Astérisque* (2008), no. 319, 1–12.

[23] C. Breuil, Série spéciale p-adique et cohomologie étale complétée. *Astérisque* (2010), no. 331, 65–115.

[24] C. Breuil, Correspondance de Langlands p-adique, compatibilité local-global et applications [d'après Colmez, Emerton, Kisin, ...]. *Astérisque* (2012), no. 348, Exp. No. 1031, viii, 119–147.

[25] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbf{Q}$: wild 3-adic exercises. *J. Amer. Math. Soc.* **14** (2001), no. 4, 843–939.

[26] C. Breuil and A. Mézard, Multiplicités modulaires et représentations de $\mathrm{GL}_2(\mathbf{Z}_p)$ et de $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ en $l = p$. *Duke Math. J.* **115** (2002), no. 2, 205–310.
[27] S. Brochard, Proof of de Smit's conjecture: a freeness criterion. *Compos. Math.* **153** (2017), no. 11, 2310–2317.
[28] W. Bruns and J. Herzog, *Cohen-Macaulay rings.* Cambridge Studies in Advanced Mathematics 39, Cambridge University Press, Cambridge, 1993.
[29] J. P. Buhler, *Icosahedral Galois representations.* Lecture Notes in Mathematics, Vol. 654, Springer-Verlag, Berlin-New York, 1978.
[30] K. Buzzard, Potential modularity—a survey. In *Non-abelian fundamental groups and Iwasawa theory*, pp. 188–211, London Math. Soc. Lecture Note Ser. 393, Cambridge Univ. Press, Cambridge, 2012.
[31] K. Buzzard, M. Dickinson, N. Shepherd-Barron, and R. Taylor, On icosahedral Artin representations. *Duke Math. J.* **109** (2001), no. 2, 283–318.
[32] K. Buzzard and T. Gee, The conjectural connections between automorphic representations and Galois representations. In *Automorphic forms and Galois representations. Vol. 1*, pp. 135–187, London Math. Soc. Lecture Note Ser. 414, Cambridge Univ. Press, Cambridge, 2014.
[33] K. Buzzard and R. Taylor, Companion forms and weight one forms. *Ann. of Math. (2)* **149** (1999), no. 3, 905–919.
[34] F. Calegari, Even Galois representations and the Fontaine-Mazur conjecture. *Invent. Math.* **185** (2011), no. 1, 1–16.
[35] F. Calegari, Even Galois Representations and the Fontaine-Mazur conjecture II. *J. Amer. Math. Soc.* **25** (2012), no. 2, 533–554.
[36] F. Calegari and N. M. Dunfield, Automorphic forms and rational homology 3-spheres. *Geom. Topol.* **10** (2006), 295–329.
[37] F. Calegari and M. Emerton, Completed cohomology—a survey. In *Non-abelian fundamental groups and Iwasawa theory*, pp. 239–257, London Math. Soc. Lecture Note Ser. 393, Cambridge Univ. Press, Cambridge, 2012.
[38] F. Calegari and D. Geraghty, Modularity lifting beyond the Taylor–Wiles method. *Inventiones mathematicae* **211** (2018), no. 1, 297–433.
[39] F. Calegari and D. Geraghty, Minimal modularity lifting for nonregular symplectic representations. *Duke Math. J.* **169** (2020), no. 5, 801–896.
[40] F. Calegari and B. Mazur, Nearly ordinary Galois deformations over arbitrary number fields. *J. Inst. Math. Jussieu* **8** (2009), no. 1, 99–177.
[41] F. Calegari and A. Venkatesh, A torsion Jacquet-Langlands correspondence. *Astérisque* (2019), no. 409, x+226.
[42] A. Caraiani, Local-global compatibility and the action of monodromy on nearby cycles. *Duke Math. J.* **161** (2012), no. 12, 2311–2413.
[43] A. Caraiani, Monodromy and local-global compatibility for $l = p$. *Algebra Number Theory* **8** (2014), no. 7, 1597–1646.
[44] A. Caraiani and B. V. Le Hung, On the image of complex conjugation in certain Galois representations. *Compos. Math.* **152** (2016), no. 7, 1476–1488.
[45] A. Caraiani and P. Scholze, On the generic part of the cohomology of compact unitary Shimura varieties. *Ann. of Math. (2)* **186** (2017), no. 3, 649–766.
[46] A. Caraiani and P. Scholze, On the generic part of the cohomology of non-compact unitary Shimura varieties. 2019, `arXiv:1909.01898`.
[47] H. Carayol, Sur les représentations l-adiques associées aux formes modulaires de Hilbert. *Ann. Sci. École Norm. Sup. (4)* **19** (1986), no. 3, 409–468.
[48] L. Clozel, Représentations galoisiennes associées aux représentations automorphes autoduales de $\mathrm{GL}(n)$. *Inst. Hautes Études Sci. Publl. Math.* (1991), no. 73, 97–145.

[49] L. Clozel, M. Harris, J.-P. Labesse, and B.-C. Ngô (eds.), *On the stabilization of the trace formula*. Stabilization of the Trace Formula, Shimura Varieties, and Arithmetic Applications, Volume I: Stabilization of the trace formula, International Press, Somerville, MA, 2011.

[50] L. Clozel, M. Harris, and R. Taylor, Automorphy for some l-adic lifts of automorphic mod l Galois representations. *Inst. Hautes Études Sci. Publ. Math.* **108** (2008), 1–181.

[51] L. Clozel and J. A. Thorne, Level-raising and symmetric power functoriality, I. *Compos. Math.* **150** (2014), no. 5, 729–748.

[52] L. Clozel and J. A. Thorne, Level raising and symmetric power functoriality, II. *Ann. of Math. (2)* **181** (2015), no. 1, 303–359.

[53] L. Clozel and J. A. Thorne, Level-raising and symmetric power functoriality, III. *Duke Math. J.* **166** (2017), no. 2, 325–402.

[54] R. Coleman and B. Mazur, The eigencurve. In *Galois representations in arithmetic algebraic geometry (Durham, 1996)*, pp. 1–113, London Math. Soc. Lecture Note Ser. 254, Cambridge Univ. Press, Cambridge, 1998.

[55] R. F. Coleman, Classical and overconvergent modular forms. *Invent. Math.* **124** (1996), no. 1-3, 215–241.

[56] P. Colmez, Les conjectures de monodromie p-adiques. 2003, séminaire Bourbaki. Vol. 2001/2002.

[57] P. Colmez, Une correspondance de langlands locale p-adique pour les représentations semi-stables de dimension 2, 2004, URL `https://webusers.imj-prg.fr/~pierre.colmez/sst.pdf`, unpublished.

[58] P. Colmez, Représentations triangulines de dimension 2. *Astérisque* (2008), no. 319, 213–258.

[59] P. Colmez, La série principale unitaire de $\mathrm{GL}_2(\mathbf{Q}_p)$. *Astérisque* (2010), no. 330, 213–262.

[60] P. Colmez, Représentations de $\mathrm{GL}_2(\mathbf{Q}_p)$ et (ϕ, Γ)-modules. *Astérisque* (2010), no. 330, 281–509.

[61] P. Colmez (ed.), *Correspondance Serre-Tate. Vol. I.* Documents Mathématiques (Paris) [Mathematical Documents (Paris)] 13, Société Mathématique de France, Paris, 2015.

[62] P. Colmez (ed.), *Correspondance Serre-Tate. Vol. II.* Documents Mathématiques (Paris) [Mathematical Documents (Paris)] 14, Société Mathématique de France, Paris, 2015.

[63] P. Colmez, G. Dospinescu, and V. Paškūnas, The p-adic local Langlands correspondence for $GL_2(\mathbb{Q}_p)$. *Camb. J. Math.* **2** (2014), no. 1, 1–47.

[64] B. Conrad, Finite group schemes over bases with low ramification. *Compositio Math.* **119** (1999), no. 3, 239–320.

[65] B. Conrad, Ramified deformation problems. *Duke Math. J.* **97** (1999), no. 3, 439–513.

[66] B. Conrad, F. Diamond, and R. Taylor, Modularity of certain potentially Barsotti-Tate Galois representations. *J. Amer. Math. Soc.* **12** (1999), no. 2, 521–567.

[67] G. Cornell, J. H. Silverman, and G. Stevens (eds.), *Modular forms and Fermat's last theorem.* Springer-Verlag, New York, 1997.

[68] H. Darmon, F. Diamond, and R. Taylor, Fermat's last theorem. In *Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993)*, pp. 2–140, Int. Press, Cambridge, MA, 1997.

[69] P. Deligne and J.-P. Serre, Formes modulaires de poids 1. *Ann. Sci. École Norm. Sup. (4)* **7** (1974), 507–530 (1975).

[70] F. Diamond, On deformation rings and Hecke rings. *Ann. of Math. (2)* **144** (1996), no. 1, 137–166.

[71] F. Diamond, An extension of Wiles' results. In *Modular forms and Fermat's last theorem (Boston, MA, 1995)*, pp. 475–489, Springer, New York, 1997.

[72] F. Diamond, The Taylor-Wiles construction and multiplicity one. *Invent. Math.* **128** (1997), no. 2, 379–391.

[73] M. Dickinson, On the modularity of certain 2-adic Galois representations. *Duke Math. J.* **109** (2001), no. 2, 319–382.

[74] L. Dieulefait, Automorphy of $\mathrm{Symm}^5(\mathrm{GL}(2))$ and base change. *J. Math. Pures Appl. (9)* **104** (2015), no. 4, 619–656.

[75] J. Elstrodt, F. Grunewald, and J. Mennicke, PSL(2) over imaginary quadratic integers. In *Arithmetic Conference (Metz, 1981)*, pp. 43–60, Astérisque 94, Soc. Math. France, Paris, 1982.

[76] M. Emerton, A local-global compatibility conjecture in the p-adic Langlands programme for $\mathrm{GL}_{2/\mathbb{Q}}$. *Pure Appl. Math. Q.* **2** (2006), no. 2, Special Issue: In honor of John H. Coates. Part 2, 279–393.

[77] M. Emerton, On the interpolation of systems of eigenvalues attached to automorphic Hecke eigenforms. *Invent. Math.* **164** (2006), no. 1, 1–84.

[78] M. Emerton, Local-global compatibility in the p-adic Langlands programme for $\mathrm{GL}_2/\mathbb{Q}$, 2010, preprint.

[79] N. Fakhruddin, C. Khare, and S. Patrikis, Lifting and automorphy of reducible mod p galois representations over global fields. 2020, `arXiv:2008.12593`.

[80] J.-M. Fontaine, Groupes finis commutatifs sur les vecteurs de Witt. *C. R. Acad. Sci. Paris Sér. A-B* **280** (1975), Ai, A1423–A1425.

[81] J.-M. Fontaine, Semi-stable Galois representations, 1988, URL `https://webusers.imj-prg.fr/~pierre.colmez/arbeitstagung-1988.pdf`, notes from the 1988 Mathematische Arbeitstagung.

[82] J.-M. Fontaine and G. Laffaille, Construction de représentations p-adiques. *Ann. Sci. École Norm. Sup. (4)* **15** (1982), no. 4, 547–608 (1983).

[83] J.-M. Fontaine and B. Mazur, Geometric Galois representations. In *Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993)*, pp. 41–78, Ser. Number Theory, I, Int. Press, Cambridge, MA, 1995.

[84] N. Freitas, B. V. Le Hung, and S. Siksek, Elliptic curves over real quadratic fields are modular. *Invent. Math.* **201** (2015), no. 1, 159–206.

[85] K. Fujiwara, Deformation rings and Hecke algebras in the totally real case. 2006, `arXiv:math/0602606`.

[86] S. Galatius and A. Venkatesh, Derived Galois deformation rings. *Adv. Math.* **327** (2018), 470–623.

[87] T. Gee, Automorphic lifts of prescribed types. *Math. Ann.* **350** (2011), no. 1, 107–144.

[88] A. Genestier and J. Tilouine, Systèmes de Taylor-Wiles pour GSp(4). *Astérisque* (2005), no. 302, 177–290.

[89] D. Geraghty, Modularity lifting theorems for ordinary Galois representations. *Math. Ann.* **373** (2019), no. 3-4, 1341–1427.

[90] W. Goldring and J.-S. Koskivirta, Strata Hasse invariants, Hecke algebras and Galois representations. *Invent. Math.* **217** (2019), no. 3, 887–984.

[91] E. S. Golod and I. R. Šafarevič, On the class field tower. *Izv. Akad. Nauk SSSR Ser. Mat.* **28** (1964), 261–272.

[92] B. Green, F. Pop, and P. Roquette, On Rumely's local-global principle. *Jahresber. Deutsch. Math.-Verein.* **97** (1995), no. 2, 43–74.

[93] T. Haines and M. Harris (eds.), *Stabilization of the trace formula, Shimura varieties, and arithmetic applications, volume II: Shimura varieties.* London Mathematical Society Lecture Note Series, Cambridge University Press, 2020.

[94] D. Hansen, Minimal modularity lifting for GL_2 over an arbitrary number field. 2012, `arXiv:1209.5309`.

[95] M. Harris, Potential automorphy of odd-dimensional symmetric powers of elliptic curves and applications. In *Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. II*, pp. 1–21, Progr. Math. 270, Birkhäuser Boston, Boston, MA, 2009.

[96] M. Harris, K.-W. Lan, R. Taylor, and J. Thorne, On the rigid cohomology of certain Shimura varieties. *Res. Math. Sci.* **3** (2016), 3:37.

[97] M. Harris, N. Shepherd-Barron, and R. Taylor, A family of Calabi-Yau varieties and potential automorphy. *Ann. of Math. (2)* **171** (2010), no. 2, 779–813.

[98] M. Harris, D. Soudry, and R. Taylor, l-adic representations associated to modular forms over imaginary quadratic fields. I. Lifting to $GSp_4(\mathbf{Q})$. *Invent. Math.* **112** (1993), no. 2, 377–411.

[99] M. Harris and R. Taylor, Deformations of automorphic Galois representations, 1998–2003, http://www.math.columbia.edu/ harris/website/publications.
[100] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties.* Annals of Mathematics Studies 151, Princeton University Press, Princeton, NJ, 2001.
[101] H. Hida, Galois representations into $\mathrm{GL}_2(\mathbf{Z}_p[[X]])$ attached to ordinary cusp forms. *Invent. Math.* **85** (1986), no. 3, 545–613.
[102] H. Hida, Nearly ordinary Hecke algebras and Galois representations of several variables. In *Algebraic analysis, geometry, and number theory (Baltimore, MD, 1988)*, pp. 115–134, Johns Hopkins Univ. Press, Baltimore, MD, 1989.
[103] H. Hida, p-adic ordinary Hecke algebras for GL(2). *Ann. Inst. Fourier (Grenoble)* **44** (1994), no. 5, 1289–1322.
[104] Y. Hu and V. Paškūnas, On crystabelline deformation rings of $\mathrm{G}al(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. *Math. Ann.* **373** (2019), no. 1-2, 421–487.
[105] H. Jacquet and J. A. Shalika, On Euler products and the classification of automorphic forms. II. *Amer. J. Math.* **103** (1981), no. 4, 777–815.
[106] P. L. Kassaei, A gluing lemma and overconvergent modular forms. *Duke Math. J.* **132** (2006), no. 3, 509–529.
[107] P. L. Kassaei, Modularity lifting in parallel weight one. *J. Amer. Math. Soc.* **26** (2013), no. 1, 199–225.
[108] P. L. Kassaei, Analytic continuation of overconvergent Hilbert modular forms. *Astérisque* (2016), no. 382, 1–48.
[109] P. L. Kassaei, S. Sasaki, and Y. Tian, Modularity lifting results in parallel weight one and applications to the Artin conjecture: the tamely ramified case. *Forum Math. Sigma* **2** (2014), e18, 58.
[110] C. Khare, Remarks on mod p forms of weight one. *Internat. Math. Res. Notices* (1997), no. 3, 127–133.
[111] C. Khare, Serre's modularity conjecture: the level one case. *Duke Math. J.* **134** (2006), no. 3, 557–589.
[112] C. Khare and J.-P. Wintenberger, On Serre's conjecture for 2-dimensional mod p representations of $\mathrm{G}al(\overline{\mathbb{Q}}/\mathbb{Q})$. *Ann. of Math. (2)* **169** (2009), no. 1, 229–253.
[113] C. Khare and J.-P. Wintenberger, Serre's modularity conjecture. I. *Invent. Math.* **178** (2009), no. 3, 485–504.
[114] C. Khare and J.-P. Wintenberger, Serre's modularity conjecture. II. *Invent. Math.* **178** (2009), no. 3, 505–586.
[115] C. B. Khare and J. A. Thorne, Potential automorphy and the Leopoldt conjecture. *Amer. J. Math.* **139** (2017), no. 5, 1205–1273.
[116] M. Kisin, Overconvergent modular forms and the Fontaine-Mazur conjecture. *Invent. Math.* **153** (2003), no. 2, 373–454.
[117] M. Kisin, Modularity for some geometric Galois representations. In *L-functions and Galois representations*, pp. 438–470, London Math. Soc. Lecture Note Ser. 320, Cambridge Univ. Press, Cambridge, 2007.
[118] M. Kisin, Modularity of 2-dimensional Galois representations. In *Current developments in mathematics, 2005*, pp. 191–230, Int. Press, Somerville, MA, 2007.
[119] M. Kisin, Potentially semi-stable deformation rings. *J. Amer. Math. Soc.* **21** (2008), no. 2, 513–546.
[120] M. Kisin, The Fontaine-Mazur conjecture for GL_2. *J. Amer. Math. Soc.* **22** (2009), no. 3, 641–690.
[121] M. Kisin, Modularity of 2-adic Barsotti-Tate representations. *Invent. Math.* **178** (2009), no. 3, 587–634.
[122] M. Kisin, Moduli of finite flat group schemes, and modularity. *Annals of Math.(2)* **170** (2009), no. 3, 1085–1180.

[123] M. Kisin, The structure of potentially semi-stable deformation rings. In *Proceedings of the International Congress of Mathematicians. Volume II*, pp. 294–311, Hindustan Book Agency, New Delhi, 2010.
[124] S. L. Kleiman, The standard conjectures. In *Motives (Seattle, WA, 1991)*, pp. 3–20, Proc. Sympos. Pure Math. 55, Amer. Math. Soc., Providence, RI, 1994.
[125] R. E. Kottwitz, On the λ-adic representations associated to some simple Shimura varieties. *Invent. Math.* **108** (1992), no. 3, 653–665.
[126] K.-W. Lan and J. Suh, Vanishing theorems for torsion automorphic sheaves on general PEL-type Shimura varieties. *Adv. Math.* **242** (2013), 228–286.
[127] R. P. Langlands, *Base change for* GL(2). Annals of Mathematics Studies 96, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980.
[128] G. Laumon and B. C. Ngô, Le lemme fondamental pour les groupes unitaires. *Ann. of Math. (2)* **168** (2008), no. 2, 477–573.
[129] B. Mazur, Deforming Galois representations. In *Galois groups over* $\mathbf{Q}$ *(Berkeley, CA, 1987)*, pp. 385–437, Math. Sci. Res. Inst. Publ. 16, Springer, New York, 1989.
[130] B. Mazur and A. Wiles, On p-adic analytic families of Galois representations. *Compositio Math.* **59** (1986), no. 2, 231–264.
[131] J. S. Milne, *Arithmetic duality theorems.* Perspectives in Mathematics 1, Academic Press, Inc., Boston, MA, 1986.
[132] L. Moret-Bailly, Groupes de Picard et problèmes de Skolem II. *Ann. Scient. Ec. Norm. Sup.* **22** (1989), 181–194.
[133] J. Newton and J. A. Thorne, Adjoint selmer groups of automorphic Galois representations of unitary type. 2019, `arXiv:1912.11265`.
[134] J. Newton and J. A. Thorne, Symmetric power functoriality for holomorphic modular forms. 2019, `arXiv:1912.11261`.
[135] J. Newton and J. A. Thorne, Symmetric power functoriality for holomorphic modular forms, ii. 2020, `arXiv:2009.07180`.
[136] B. C. Ngô, Le lemme fondamental pour les algèbres de Lie. *Publ. Math. Inst. Hautes Études Sci.* (2010), no. 111, 1–169.
[137] L. Pan, The Fontaine–Mazur conjecture in the residually reducible case. 2019, `arXiv:1901.07166`.
[138] L. Pan, On locally analytic vectors of the completed cohomology of modular curves. 2020, `arXiv:2008.07099`.
[139] S. Patrikis and R. Taylor, Automorphy and irreducibility of some l-adic representations. *Compos. Math.* **151** (2015), no. 2, 207–229.
[140] V. Paškūnas, The image of Colmez's Montreal functor. *Publ. Math. Inst. Hautes Études Sci.* **118** (2013), 1–191.
[141] V. Pilloni, Formes modulaires p-adiques de Hilbert de poids 1. *Invent. Math.* **208** (2017), no. 2, 633–676.
[142] V. Pilloni, Higher coherent cohomology and p-adic modular forms of singular weights. *Duke Math. J.* **169** (2020), no. 9, 1647–1807.
[143] V. Pilloni and B. Stroh, Surconvergence, ramification et modularité. *Astérisque* (2016), no. 382, 195–266.
[144] L. Qian, Potential automorphy for GL_n. 2021, `arXiv:2104.09761`.
[145] R. Ramakrishna, On a variation of Mazur's deformation functor. *Compositio Math.* **87** (1993), no. 3, 269–286.
[146] R. Ramakrishna, Lifting Galois representations. *Invent. Math.* **138** (1999), no. 3, 537–562.
[147] R. Ramakrishna, Deforming Galois representations and the conjectures of Serre and Fontaine-Mazur. *Ann. of Math. (2)* **156** (2002), no. 1, 115–154.

[148] M. Raynaud, *Théorèmes de Lefschetz en cohomologie cohérente et en cohomologie étale*. Supplément au Bull. Soc. Math. France, Tome 103, Société Mathématique de France, Paris, 1975.
[149] K. A. Ribet, On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms. *Invent. Math.* **100** (1990), no. 2, 431–476.
[150] K. A. Ribet, Galois representations and modular forms. *Bull. Amer. Math. Soc. (N.S.)* **32** (1995), no. 4, 375–402.
[151] T. Saito, Modular forms and p-adic Hodge theory. *Invent. Math.* **129** (1997), no. 3, 607–620.
[152] S. Sasaki, Integral models of Hilbert modular varieties in the ramified case, deformations of modular Galois representations, and weight one forms. *Invent. Math.* **215** (2019), no. 1, 171–264.
[153] D. Savitt, Modularity of some potentially Barsotti-Tate Galois representations. *Compos. Math.* **140** (2004), no. 1, 31–63.
[154] D. Savitt, On a conjecture of Conrad, Diamond, and Taylor. *Duke Math. J.* **128** (2005), no. 1, 141–197.
[155] P. Scholze, On torsion in the cohomology of locally symmetric varieties. *Ann. of Math. (2)* **182** (2015), no. 3, 945–1066.
[156] J.-P. Serre, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. J.* **54** (1987), no. 1, 179–230.
[157] N. Shepherd-Barron and R. Taylor, On degree 2 Galois representations over $\mathbf{F}_4$. *Proc. Nat. Acad. Sci. U.S.A.* **94** (1997), no. 21, 11147–11148.
[158] S. W. Shin, Galois representations arising from some compact Shimura varieties. *Ann. of Math. (2)* **173** (2011), no. 3, 1645–1741.
[159] C. M. Skinner and A. J. Wiles, Ordinary representations and modular forms. *Proc. Nat. Acad. Sci. U.S.A.* **94** (1997), no. 20, 10520–10527.
[160] C. M. Skinner and A. J. Wiles, Residually reducible representations and modular forms. *Inst. Hautes Études Sci. Publ. Math.* (1999), no. 89, 5–126 (2000).
[161] C. M. Skinner and A. J. Wiles, Base change and a problem of Serre. *Duke Math. J.* **107** (2001), no. 1, 15–25.
[162] C. M. Skinner and A. J. Wiles, Nearly ordinary deformations of irreducible residual representations. *Ann. Fac. Sci. Toulouse Math. (6)* **10** (2001), no. 1, 185–215.
[163] A. Snowden, Singularities of ordinary deformation rings. *Math. Z.* **288** (2018), no. 3-4, 759–781.
[164] J. Tate, The non-existence of certain Galois extensions of $\mathbf{Q}$ unramified outside 2. In *Arithmetic geometry (Tempe, AZ, 1993)*, pp. 153–156, Contemp. Math. 174, Amer. Math. Soc., Providence, RI, 1994.
[165] R. Taylor, l-adic representations associated to modular forms over imaginary quadratic fields. II. *Invent. Math.* **116** (1994), no. 1-3, 619–643.
[166] R. Taylor, Remarks on a conjecture of Fontaine and Mazur. *J. Inst. Math. Jussieu* **1** (2002), no. 1, 125–143.
[167] R. Taylor, Galois representations. *Ann. Fac. Sci. Toulouse Math. (6)* **13** (2004), no. 1, 73–119.
[168] R. Taylor, The image of complex conjugation in l-adic representations associated to automorphic forms. *Algebra Number Theory* **6** (2012), no. 3, 405–435.
[169] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)* **141** (1995), no. 3, 553–572.
[170] R. Taylor and T. Yoshida, Compatibility of local and global Langlands correspondences. *J. Amer. Math. Soc.* **20** (2007), no. 2, 467–493.
[171] R. L. Taylor, *On congruences between modular forms*. ProQuest LLC, Ann Arbor, MI, 1988.
[172] J. A. Thorne, On the automorphy of l-adic Galois representations with small residual image. *J. Inst. Math. Jussieu* **11** (2012), no. 4, 855–920.
[173] J. A. Thorne, Automorphy lifting for residually reducible l-adic Galois representations. *J. Amer. Math. Soc.* **28** (2015), no. 3, 785–870.

[174] J. Tilouine and E. Urban, Integral period relations and congruences. 2018, arXiv:1811.11166.
[175] T. Tsuji, p-adic étale cohomology and crystalline cohomology in the semi-stable reduction case. *Invent. Math.* **137** (1999), no. 2, 233–411.
[176] S.-N. Tung, On the modularity of 2-adic potentially semi-stable deformation rings. *Math. Z.* **298** (2021), no. 1-2, 107–159.
[177] J. Tunnell, Artin's conjecture for representations of octahedral type. *Bull. Amer. Math. Soc. (N.S.)* **5** (1981), no. 2, 173–175.
[178] G. Wiese, On Galois representations of weight one. *Doc. Math.* **19** (2014), 689–707.
[179] A. Wiles, On ordinary λ-adic representations associated to modular forms. *Invent. Math.* **94** (1988), no. 3, 529–573.
[180] A. Wiles, Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)* **141** (1995), no. 3, 443–551.
[181] X. Zhu, Coherent sheaves on the stack of Langlands parameters. 2021, arXiv:2008.02998.

13. Acknowledgments

The arithmetic Langlands program has been a bustling field of activity in the past 30 years with many deep results by a number of extraordinary mathematicians. To be part of this field (mostly as an observer but occasionally as a contributor) has been an amazing experience. I would like to thank my collaborators all of whom have taught me so much, with special thanks to Matthew Emerton and Toby Gee who have in addition been my close friends. Thanks to Toby Gee and Ravi Ramakrishna for a number of useful suggestions and corrections on the first version of this survey. Finally, special thanks to those mathematicians who allowed me to share some of their personal recollections in this paper, in particular Christophe Breuil, Ana Caraiani, Pierre Colmez, Fred Diamond, Michael Harris, Mark Kisin, Peter Scholze, Richard Taylor, and Andrew Wiles.

The University of Chicago, 5734 S University Ave, Chicago, IL 60637, USA fcale@uchicago.edu