Galois representations and torsion cohomology

A series of misunderstandings

Columbia University

February 1, 2021

Outline

Locally symmetric spaces

We will look at $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ or $Gal(\overline{\mathbb{Q}}/K)$, where $K = \mathbb{Q}(\sqrt{-d})$ is an imaginary quadratic field.
But everything is known (or could be known) whenever K is either a *totally real number field* or a *CM field* (totally imaginary quadratic extension of a totally real field). Complete understanding is only possible in this generality.

Locally symmetric spaces

We consider the locally symmetric spaces attached to $GL(n)$ over $\mathbb{Q}$ or K:

$$X_\Gamma := \Gamma \backslash GL(n, K \otimes \mathbb{R}) / K_\infty \cdot Z(\mathbb{R})$$

where $\Gamma \subset GL(n, \mathcal{O}_K)$ is a (congruence) subgroup of finite index, $Z(\mathbb{R})$ is the diagonal subgroup of $GL(n, K \otimes \mathbb{R})$, and K_∞ is a maximal compact subgroup: either $SO(n)$ (over $\mathbb{Q}$) or $U(n)$ (over K).
We actually have to consider all Γ simultaneously (the adelic locally symmetric space) but I don't want to write the definition; so I just write the notation: $X_{n,\mathbb{Q}}$ or $X_{n,K}$.

A theorem about cohomology

This space has cohomology and we start with

$$H^*_!(X_{n,\mathbb{Q}}, \mathbb{Q}) := Im[H^*_c(X_{n,\mathbb{Q}}, \mathbb{Q}) \to H^*(X_{n,\mathbb{Q}}, \mathbb{Q})].$$

(likewise with $X_{n,K}$)
This space is (in some sense) finite-dimensional (depending on Γ, in some sense) and has a large commuting $\mathbb{Q}$-algebra $\mathbb{T}$ of operators.

Theorem

Let $\alpha : \mathbb{T} \to \mathbb{Q}$ be a homomorphism. Then for any prime number ℓ, there is an n-dimensional representation

$$\rho_{\alpha,\ell} : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL(n, \mathbb{Q}_\ell) \;\; (\rho_{\alpha,\ell} : Gal(\overline{\mathbb{Q}}/K) \to GL(n, \mathbb{Q}_\ell))$$

unramified outside a finite set of primes S, such that, for all $p \notin S$ the characteristic polynomial of Frobenius at p is determined by the values of α on the Hecke operators in $\mathbb{T}$ corresponding to p.

Clozel's theorem

When $n = 2$ this was proved (for $\mathbb{Q}$) by Eichler and Shimura, and this is the starting point for the proof of Fermat's Last Theorem.
When $n = 3$ this is a UCLA theorem, essentially due to Blasius and Rogawski (worked out in 1988 in Montreal).

Theorem (Clozel)

Suppose α is contained in the self-dual part of the cohomology. (The space $H^*_!(X_{n,\mathbb{Q}}, \mathbb{Q})$ satisfies Poincaré duality and we consider a $\mathbb{T}$-eigenspace that is its own dual; or complex-conjugate to its dual for K.) Suppose S is not empty (and some additional hypotheses that were eventually relaxed). Then the $\rho_{\alpha,\ell}$ exist.

Remark

We can take $\alpha : \mathbb{T} \to E$ for any number field E, and the theorem is still valid.

Clozel's theorem announced in Ann Arbor, 1988

(A) Used the *stable trace formula* to relate α to the cohomology of a *Shimura variety* $S_{n,K}$ obtained as an arithmetic quotient of the unit ball in $\mathbb{C}^{n-1}$, with a *canonical model* over the (imaginary quadratic field) K, so its topological cohomology can be related to its ℓ-adic cohomology, which has a Galois action that commutes with appropriate Hecke operators.
(B) The $Gal(\overline{\mathbb{Q}}/K)$-action on this cohomology had just been determined by Kottwitz (also announced in Ann Arbor).
(C) Q.E.D.

The idea almost announced in Ann Arbor

Laurent Clozel

This theorem came as a great surprise to me, because at the time the case $n = 3$ was still being written up. I spoke to Clozel after his talk and he explained that in his next talk he was going to announce his plans to solve the remaining (not self-dual) case.

The space $X_{n,K}$ is part of the (Borel-Serre) boundary of a Shimura variety $Y_{2n,K}$ related to the Lie group $U(n,n)$ (unitary group in $2n$-variables over K, with signature (n,n)).

The α-eigenspace of the cohomology of $X_{n,K}$ could then be realized as an eigenspace in the cohomology of $Y_{2n,K}$.

But this eigenspace (over $\mathbb{Q}_\ell$) also has a Galois action. What could it be if not the one predicted by the Langlands conjectures?

Right cohomology, wrong Galois action

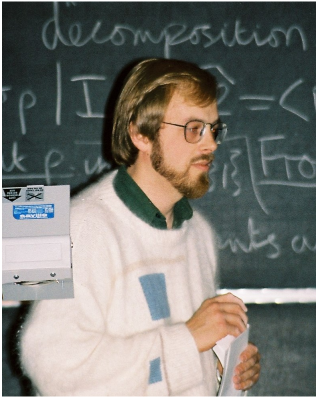
Here is where my role in the story begins, primarily as an engaged spectator. I had been thinking about the boundary cohomology of Shimura varieties in terms of the *toroidal compactification*. The boundary cohomology attached to $X_{n,K}$ all comes from *rational varieties*. I explained to Clozel that Galois action looked too simple to be the one predicted by Langlands.

I worked this all out in detail over the next few years with Zucker (starting in Ann Arbor). In particular, we showed that boundary cohomology had a *weight filtration* corresponding to mixed Hodge theory.

Clozel checked with Kottwitz, who agreed with my diagnosis. So Clozel took the theorem for non self-dual representations off the agenda for the time being.

Steve Zucker

Collaboration with Taylor, $n = 2$



Richard Taylor 8 years after Ann Arbor

I met Richard Taylor for the first time in Ann Arbor. A few years later, Soudry and I provided the missing piece in his project to prove the theorem for non-self-dual representations of $GL(2)$ over K, using the theta correspondence; this was my first collaboration with Taylor.

(Our work was generalized by Chung-Pang Mok 20 years later to $GL(2)$ over any CM field.)

Collaboration with Taylor, $n > 2$

A few years after our paper with Soudry, Taylor and I refined Clozel's theorem using p-adic uniformization of Shimura varieties (at the primes in S), obtaining the local Langlands correspondence as a corollary.

The result was gradually refined over the next ten years, with important contributions by Taylor-Yoshida, Labesse, Clozel, Shin, Chenevier, and Caraiani, leading to the removal of successive ramification conditions. (Removal of the final ramification conditions will be mentioned a few slides from now.)

Skinner suggests using congruences

Chris Skinner, undated photo

I met Chris Skinner in 2000. At some point after we met, probably in 2002, he returned to Clozel's idea. He had written a paper with Eric Urban on eigenvarieties, and he suggested that the boundary cohomology of $Y_{2n,K}$, coming from $X_{n,K}$, could be deformed p-adically to classes in the interior of $Y_{2n,K}$.

These classes had the right kind of $2n$-*dimensional* Galois representation to be split up into two n-dimensional representations, one of which was the one predicted by the Langlands correspondence (he had checked).

Skinner suggests using congruences

I filed this away in my mind: Chris Skinner, possibly in collaboration with Urban, was going to use eigenvarieties to prove the non self-dual case of the theorem I stated.

I was happy for them although I knew it would take a long time, because they were busy writing up their proof of the main conjecture for the p-adic L-functions of elliptic curves, using (of course) eigenvarieties.

Chris Skinner, undated photo

The Montreal conference, September 2005

I arrived at the conference in Montreal intending to talk about the **Paris book project** (after climbing the stairs on my knees).

Laumon and Ngô had proved the Fundamental Lemma for unitary groups (soon Ngô would prove it in general) and the book project would work to remove the ramification conditions that persisted in my book with Taylor, using the full stable trace formula.

In the middle of my talk I unexpectedly announced that Chenevier's work with Bellaïche on eigenvarieties would provide the final step although I only realized several weeks later that this is what I had done.

Breakfast in Montreal, 2005

Barry Mazur and friend

However, I was still on Paris time, and I came down very early for breakfast every day. Barry Mazur was staying at the same hotel and he is an early riser. So over breakfast, he explained to me something that was on his mind. He had long been convinced that most of the Galois representations attached to the cohomology of $GL(n, \mathbb{Q})$ for $n > 2$, or $GL(n, K)$ for $n \geq 2$, came from *torsion* classes. His recent paper with Calegari had confirmed this in the first non-trivial case.

Breakfast in Montreal, 2005

The trace formula methods of the Paris book project knew nothing about torsion cohomology. But Mazur was convinced that there had to be Galois representations attached to torsion classes. I later learned that Serre and Taylor, and Ash and Stevens, among others, had also come to this conclusion. I filed this away in my mind as a mystery.

I assumed Skinner and Urban would eventually work out the non self-dual case. But the idea that Galois representations could be attached to torsion classes seemed to me so far-fetched that I could not get it out of my mind.

A uniquely satisfying idea

At some point in 2006, I realized that my work with Zucker showed that torsion classes in $H^*(X_{n,K})$ could also be realized in $H^*(Y_{2n,K})$.
This was the first and (to all intents and purposes) the last idea I had in connection with this question. But it was uniquely satisfying for three reasons.

- It suggested an unexpected application of my work with Zucker.
- My work on the Sato-Tate conjecture (with Clozel, Shepherd-Barron, and Taylor) had just come to a successful conclusion, and I was looking for a new problem.
- It reminded me of my pleasant breakfast conversations with Barry Mazur in Montreal. My main motivation in giving this talk is to stress how important that experience was for everything that followed. Such considerations are unfortunately rarely preserved in the published record.

Weights

The idea, then, was to apply Skinner's suggestion to deform *torsion* cohomology classes of $X_{n,K}$ p-adically to interior classes of $Y_{2n,K}$. For large p we could even hope to use the mixed Hodge weights to lift torsion boundary classes to the cohomology of the Shimura variety (as in my work with Zucker).

I talked about it mainly with Skinner, but also with Urban, Calegari, and Emerton. I mentioned it to Taylor in passing. I hoped that, even at the cost of imposing highly restrictive hypotheses, we could construct *at least one* Galois representation attached to a torsion class.

My idea will not work

In the spring of 2007 I flew from Paris to work with Skinner in Princeton. On the second day Urban came down from New York to join us, and to explain why the idea would not work. In fact, Skinner's original idea, to deform characteristic zero cohomology classes, could not work with the known constructions of eigenvarieties (Euler characteristics are constant).

Urban proposed a more complicated construction, based on his ongoing work with Skinner. I returned to Paris fully discouraged about the project.

Eric Urban, undated photo

Completed cohomology

Frank Calegari and Matt Emerton, 2016

In 2009 I visited Calegari and Emerton in Chicago. They wanted to talk about a completely different approach, for $GL(4)$, using their conjectures on completed cohomology. Assuming their conjectures and also an extension of the Arthur conjectures to torsion classes, Galois representations could be attached to some torsion classes for $GL(4)$. This involved a specific classification problem that was not at all obvious but that was at least concrete.

It looked more hopeful at this point than the idea with Skinner and Urban. But both ideas were extremely technical and could only be applied in low dimension.

A walk in the woods

By the time I visited the IAS in the winter 2011 I was giving little thought to the torsion question, and none at all to the non self-dual representations in characteristic zero. But a few days after I arrived, Richard Taylor invited me on a walk through the frozen Institute Woods.

The Institute Woods, photo by Mark Goresky

p-adic modular forms

Few methods were available at the time for deforming topological cohomology. But the theory of overconvergent p-adic modular forms was well understood.

Taylor's idea was to use the control provided by the overconvergent theory to compute the global p-adic de Rham cohomology $H^*_{dR}(Y_{2n,K})$ on the ordinary locus. The latter is (close to) affine, so its cohomology can be computed by global sections of the de Rham complex. This gives a complex of p-adic Banach spaces but the overconvergent theory provides enough control.

The Hecke eigenvectors of $H^*_{dR}(Y_{2n,K})$ could then be approximated p-adically by eigenspaces on holomorphic cusp forms of various weights. Galois representations are attached to the latter, and because they are cuspidal these have the right properties.

The collaboration with Lan, Taylor, and Thorne

Lan and Thorne were also at the IAS, and Taylor proposed that we work this out together. This involved solving the following problems (among others):

- Replacing the ordinary locus in the toroidal compactification (not affine) by one in the minimal compactification (affine). (This argument was discovered independently by Andreatta, Iovita, and Pilloni.)
- Finding a cohomology theory that related p-adic de Rham cohomology to p-adic modular forms, and with a weight formalism.
- Showing that the boundary classes contributed non-trivially to the cohomology of the ordinary locus.
- Doing all of this for Kuga-Sato varieties, not least because the p-adic approximation involved cohomology with arbitrarily twisted coefficients.
- Relating the coherent cohomology of Kuga-Sato varieties to that of the base Shimura variety (where the Galois representations were defined).

Dagger spaces

Taylor decided that we would use *rigid cohomology* of *dagger spaces*. This had a weight formalism (Chiarellotto) and a relation to coherent cohomology of rigid analytic or dagger spaces (Le Stum, Grosse-Klönne).

Crucially, we had to look at de Rham (coherent) cohomology of the ordinary locus that was compactly supported near the boundary but with no support condition away from the ordinary locus. I don't know whether this can now be done with less esoteric theories of p-adic analytic spaces.

The contribution of the boundary classes consisted in the weight zero subspace of cohomology with compact support. This all came from the rational varieties (as I had explained to Clozel more than 20 years earlier) but was by far the simplest part of my computations with Zucker.

Division of labor

My contribution was essentially nil, beyond my (much) earlier work with Zucker.

Concretely, I was sometimes asked to explain the relevant portions of my papers with Zucker (which even Zucker and I found hard to read). I also carried out some calculations in that framework that turned out to be unnecessary for the final results and did not appear in the paper.

Most of the technical parts of the paper were written by Lan and Taylor working closely together. Lan also had to write a second 500+ page book to justify the claims about the compactification of the ordinary loci of Kuga-Sato varieties.

While my three collaborators were busy writing the paper, I was consulting all the experts I knew in p-adic Hodge theory in the hope of finding an integral structure that could replace rigid cohomology in our construction.

Making the results public

The coefficients were all in characteristic zero – no integral theory, and therefore no torsion classes, could be treated by this theory.

Another potential advantage of an integral structure was that the finite coefficients became trivial over finite covers of $Y_{2n,K}$, eliminating the need for the mass of notation needed to work with Kuga-Sato varieties.

Taylor announced the results in the spring of 2012, before the writing was complete. I decided to talk about the project that summer at Oberwolfach, where I knew I would be surrounded by specialists in p-adic cohomology.

Beer

My own talk described the method of [HLTT]. From the Oberwolfach report:

> An important observation is that the relevant Eisenstein cohomology classes can be realized geometrically in the weight 0 subspace of *rigid cohomology* of the ordinary locus of the special fiber of $\mathfrak{X}_U$, with compact supports in the direction of the toroidal boundary. This can in turn be calculated by a spectral sequence whose $E_1^{r,s}$ terms are given by coherent cohomology of automorphic vector bundles $\mathfrak{X}_U$, extended in a certain way to $\mathfrak{X}_U^{tor}$, and then to $\mathfrak{X}_U^*$. Using the fact that the ordinary locus in the special fiber of $\mathfrak{X}_U^*$ is affine, the higher coherent cohomology all vanishes, which implies that the Eisenstein classes can be approximated modulo arbitrarily high powers of p by (holomorphic) cusp forms. Standard techniques due to Taylor and others then show that the systems of Hecke eigenvalues on the Eisenstein classes are approximated in a similar way by cuspidal Hecke eigenvalues. This gives a first construction of the pseudorepresentations predicted by Skinner, and by refining this construction one obtains the desired n-dimensional p-adic Galois representation.

My most vivid memory of that summer's Oberwolfach meeting, however, was of the largest collection of bottles of beer I had ever seen on a single table, when I went to bed at around 2 AM. Somehow they had all disappeared in time for breakfast the following morning.

More beer

My Oberwolfach report ended optimistically:

> Many questions remain open; the most intriguing is whether this technique can be extended to attach Galois representations to *torsion* cohomology of the locally symmetric spaces attached to $GL(n)$.

The talk itself ended (or began?) with a prophesy: that in 5 (or 3?) years the construction would be carried out in an integral Hodge theory, with no need for the Kuga-Sato varieties. At that point I couldn't help looking at Peter Scholze, who was sitting (as usual) at the back of the room.

I next saw Scholze at the Fields Institute in Toronto that fall. Together with Matt Emerton and a few others, we went to a pub, and I asked him about his current projects. He mentioned that he knew how to construct perfectoid Shimura varieties, and described the Hodge-Tate period morphism. He also explained that he could compute completed cohomology as coherent cohomology of these perfectoid spaces.

I reminded him of my Oberwolfach prophesy and asked whether this might be the missing integral theory.

Scholze "made some progress"

Lan and I were both invited to Bonn the following spring to talk about [HLTT], which was not yet available as a preprint.

I have to confess that the prospect of spending time in Bonn has never appealed to me. However, I hoped to chat about the integral theory with Scholze, who had just defended his thesis – after having been named the youngest professor in German history. We did have a conversation in his office at the university.

I gave two talks and returned to Paris; Lan also gave a few talks. About a month later, when I saw Scholze in Paris, he told me, "I've made some progress."

A few words about Scholze's proof

The only overlap of Scholze's *Annals* paper with the strategy of [HLTT] – which was only posted on our websites a few months later, and on arXiv a year after that! – was in the recovery of the n-dimensional Galois representations attached to torsion classes from the $2n$-dimensional representations. As far as I can tell, the most difficult material in Scholze's paper had to do with the proof of the results he had mentioned in Toronto but had not yet written up. The rest of the paper was the sort of display of abundant originality that most mathematicians don't manage in an entire lifetime, and that number theorists are still trying to digest, although Scholze himself has progressed through at least three landmark contributions in the meantime.

To mention just one example, the Hodge-Tate morphism, which gave Scholze the benefits of an affine covering while sidestepping all the constructions of compactifications and related notation – and, as I anticipated, the need to introduce Kuga-Sato varieties – that took up at least 1/3 of [HLTT], looks likely to become a fundamental object of study in its own right.

A few applications

- Pilloni and Stroh very quickly proved that Scholze's construction also provided a general construction of Galois representations attached to coherent cohomology of Shimura varieties (of Hodge type, for the specialists). This extends the construction that began with Clozel to automorphic forms whose infinity type belongs to the (non-degenerate) limit of discrete series.
- Caraiani and Scholze initiated a study applying p-adic Hodge theory, in Scholze's version, to prove vanishing of torsion cohomology under rather general conditions.
- This program continues, but it had an immediate application in the landmark *ten author paper* that proves potential automorphy of elliptic curves over CM fields.
- Most recently, Lue Pan combined the Hodge-Tate morphism with considerations inspired by the p-adic Simpson correspondence, and representation theory of enveloping algebras, to give a new proof of the Artin conjecture for odd 2-dimensional representations of $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$.

Boxer, Goldring, Koskivirta

A year after Scholze's paper was released – and a year before the 300+ pages of [HLTT] were finally published – George Boxer explained an alternative construction of the Galois representations attached to torsion classes. Boxer defined higher Hasse invariants in the setting of EGA-style algebraic geometry – he even cites EGA in his (unpublished) thesis – and works with the integral structure of provided by the coherent cohomology of schemes. The construction of Galois representations has not yet appeared, as far as I know.

Boxer's ideas overlap with constructions discovered independently by Goldring and Koskivirta, who have published a complete construction of Galois representations for *coherent cohomology* along these lines, including for torsion classes. These ideas have apparently influenced the development of *higher Hida theory* by Pilloni. This is still very much in flux.

What remains of [HLTT]

As I expected, the painstaking constructions of [HLTT] have largely proved unnecessarily complicated for the purpose, though this happened rather more quickly than I predicted.

On the other hand, Lan's 500+ page book has been indispensable for applications of the Hida theory of holomorphic modular forms in higher dimensions. My paper on p-adic L-functions with Eischen, Li, and Skinner makes extensive use of Lan's book, and it would probably have been impossible to complete otherwise.

The most interesting applications of [HLTT], as of any long and difficult paper, are the ones that no one has yet anticipated.