

MODULI STACKS OF (φ, Γ) -MODULES: A SURVEY

MATTHEW EMERTON AND TOBY GEE

CONTENTS

Lecture 1.	Introduction	2
Lecture 2.	Stacks, etc.	8
Lecture 3.	Definitions related to étale φ -modules	11
Lecture 4.	Moduli stacks of φ -modules	16
Lecture 5.	Stacks of (φ, Γ) -modules	21
Lecture 6.	Crystalline and semistable moduli stacks	25
Lecture 7.	The Herr Complex, and families of extensions	32
Lecture 8.	Crystalline Lifts	39
Lecture 9.	Geometric Breuil–Mézard	44
Lecture 10.	Moduli Spaces and Bernstein Centers	48
	References	55

This survey article is based on 10 lectures that we gave at the September 2019 Hausdorff School in Bonn on “the Emerton–Gee stack and related topics”, and the sections correspond to the original lectures. We have retained the somewhat informal style of some of these lectures, but have filled them out with further details. We hope that the first 9 lectures can serve as an extended introduction to [EG19]; the reader might first wish to read the actual introduction to [EG19], which in particular motivates the results described here. We also refer to [EG19] for any definitions not given in these lectures. The 10th lecture is a speculative look ahead to connections between our stacks and the p -adic local Langlands correspondence, and in particular explains some of our work in progress with Andrea Dotto [DEG].

We would like to thank the organisers (Johannes Anschütz, Arthur–César Le Bras, and Andreas Mihatsch) and participants of the Hausdorff school for the many helpful questions and corrections resulting from our lectures. We would especially like to thank Ashwin Iyengar for transcribing ME’s lectures at the Hausdorff school, and for allowing us to use his LaTeX files as a basis for some sections of this article. Some material in Lecture 8 is based on the notes for a series of lectures that TG gave at the ENS Lyon in June 2018, and we are grateful to the organisers (Laurent Berger, Gabriel Dospinescu, Philippe Gille, Wiesława Nizioł, Vincent Pilloni, Sandra Rozensztajn, Amaury Thuillier) of the thematic trimester on “Algebraic Groups and Geometrization of the Langlands Program” for the invitation to speak there, and the encouragement to write up notes for the lectures.

TG would like to thank the organisers (Najmuddin Fakhruddin, Eknath Ghate, Arvind Nair, C. S. Rajan, and Sandeep Varma) of the International Colloquium on Arithmetic Geometry at TIFR in January 2020, for the invitation to speak at the

conference, for their hospitality, and for the invitation to submit this article to the proceedings, without which it is unlikely that we would have completed it.

All references to [EG19] are to the second arXiv version available at <https://arxiv.org/abs/1908.07185v2>, and may not agree with the numbering in any future versions.

LECTURE 1. INTRODUCTION

As the title suggests, the paper [EG19] is about moduli stacks of (φ, Γ) -modules. In fact, we're really interested in the moduli of (local) (p -adic and mod p) Galois representations, so we'll first give some background on this problem, and the complications that arise. The theory of (φ, Γ) -modules and the construction of our moduli stacks will be the topic of subsequent lectures.

1.1. Galois representations. Let p be prime, and let K/\mathbf{Q}_p be a finite extension. We will be interested in p -adic and mod p representations of the absolute Galois group $G_K := \text{Gal}(\overline{K}/K)$. In the mod p case, this means that we will fix some $d \geq 1$ and study continuous homomorphisms

$$G_K := \text{Gal}(\overline{K}/K) \rightarrow \text{GL}_d(\overline{\mathbf{F}}_p).$$

Note these factor through $\text{GL}_d(\mathbf{F}_q)$ for some q . Similarly we could look at continuous homomorphisms

$$G_K \rightarrow \text{GL}_d(\overline{\mathbf{Q}}_p);$$

these will factor through $\text{GL}_d(E)$ for some finite E/\mathbf{Q}_p , and using compactness, they can be conjugated so as to factor through $\text{GL}_d(\mathcal{O}_E)$.

We want to arrange these representations into algebraic families (this is more-or-less what it means to construct a moduli space of d -dimensional representations of G_K), but the topological nature of G_K makes this difficult to do directly. We will elaborate on these difficulties, before turning to our method of solution (which is to pass from G_K -representations to (φ, Γ) -modules).

1.1.1. Example. The first example of an ‘‘algebraic family’’ we might consider is a family of unramified characters

$$\text{ur}_x : G_K \rightarrow G_K/I_K \cong \text{Frob}^{\widehat{\mathbf{Z}}} \rightarrow \overline{\mathbf{F}}_p^\times$$

which sends Frobenius to $x \in \overline{\mathbf{F}}_p^\times$ (note this is well-defined because any element of $\overline{\mathbf{F}}_p^\times$ lives in \mathbf{F}_q^\times for some q , so we define the unramified character by $\widehat{\mathbf{Z}} \rightarrow \mathbf{Z}/(q-1)\mathbf{Z} \rightarrow \mathbf{F}_q^\times$). But this doesn't quite come together as an algebraic family: if we wanted to make this algebraic, we would let $\mathbf{F}_p[x, x^{-1}]^\times$ be our ring of coefficients and take specializations to $\overline{\mathbf{F}}_p^\times$. We can define a homomorphism $\text{Frob}^{\mathbf{Z}} \rightarrow \mathbf{F}_p[x, x^{-1}]^\times$ taking $\text{Frob} \mapsto x$, but this doesn't extend continuously to $\text{Frob}^{\widehat{\mathbf{Z}}}$. So this gives the first basic obstruction to constructing the right moduli space.

As we will explain in the rest of this section, to get around the problem in the example, you could consider the Weil group, which is a ‘‘decompletion’’ of $\text{Frob}^{\widehat{\mathbf{Z}}}$; but ultimately, this only works well when either $d = 1$ or we work with coefficient rings in which p is invertible, so when we work with p -adic coefficients in the case $d > 1$ we have to use integral p -adic Hodge theory (and specifically the theory of (φ, Γ) -modules), as in [EG19].

1.2. Weil–Deligne Case. We begin by recalling the “Weil–Deligne formalism”, which is a standard and effective approach to dealing with topological issues in the theory of ℓ -adic Galois representations of G_K , when $\ell \neq p$. As we will see, it is of more dubious merit in the p -adic context. The ideas we explain are dealt with in much more detail (and for representations into more general groups than GL_d) in [DHKM20] and [Zhu20].

Note the local Galois group G_K has a tame quotient and pro- p wild inertia subgroup:

$$P_K \hookrightarrow G_K \twoheadrightarrow G_K^{\mathrm{tame}} = (\mathbf{Z} \rtimes \mathbf{Z}[1/q])^\wedge$$

where $1 \in \mathbf{Z}$ acts by multiplication by q on $\mathbf{Z}[1/q]$, and we take the profinite completion. We can then define the “Weil–Deligne group” WD_K as usual, which fits into the following diagram:

$$\begin{array}{ccccccc} 0 & \longrightarrow & P_K & \longrightarrow & \mathrm{WD}_K & \longrightarrow & \mathbf{Z} \rtimes \mathbf{Z}[1/q] \longrightarrow 0 \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & P_K & \longrightarrow & G_K & \longrightarrow & (\mathbf{Z} \rtimes \mathbf{Z}[1/q])^\wedge \longrightarrow 0 \end{array}$$

Then for any open and finite index subgroup $Q \leq P_K$, WD_K/Q is a finitely presented discrete group, and

$$\mathrm{WD}_K = \varprojlim_Q \mathrm{WD}_K/Q.$$

So we have essentially “decompleted and discretized” the tame part, but we are remembering the topology on the wild part of the Galois group. Furthermore, $G_K/Q = (\mathrm{WD}_K/Q)^\wedge$, and since representations of WD_K/Q with values in finite rings do extend over the profinite completion, this seems like a good first approach to defining a moduli space.¹

1.2.1. Definition. Let $V_Q \rightarrow \mathrm{Spec} \mathbf{Z}$ denote the scheme parameterizing representations $\rho : \mathrm{WD}_K/Q \rightarrow \mathrm{GL}_d$.

Note that WD_K/Q is a finitely presented group, so it’s easy to find a finite presentation for V_Q as an affine scheme. If $Q' \subseteq Q$, then there is a natural closed immersion $V_Q \hookrightarrow V_{Q'}$ (this is easy to check using the moduli description), so we can study the Ind-scheme

$$V := \varinjlim_Q V_Q.$$

By construction, any continuous $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})$ (for \mathbf{F} any finite field) factors through a finite quotient of G_K , and thus a finite quotient of WD_K , and thus factors through one of the WD_K/Q , so gives an \mathbf{F} -valued point of V . Conversely a representation $\mathrm{WD}_K/Q \rightarrow \mathrm{GL}_d(\mathbf{F})$ extends continuously to G_K/Q and thus to G_K . Therefore

$$V(\mathbf{F}) = \{\text{continuous } \bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})\}.$$

¹One technical point is that this definition of WD_K , and thus the consequent definition of a moduli space, depend on choices: namely, a choice of a lift of Frobenius and also a choice of a generator of tame inertia. We don’t dwell on this issue here, but note that Scholze and Zhu have explained how to make the same construction more canonically (involving no auxiliary choices) in the case of ℓ -adic representations when $\ell \neq p$ [Zhu20, Props. 3.1.10, 3.1.11].

If $R_{\bar{\rho}}^{\square}$ denotes the universal lifting ring of $\bar{\rho}$ one can check that there is a natural map

$$\mathrm{Spf} R_{\bar{\rho}}^{\square} \rightarrow V.$$

In fact, this map is versal to V at $\bar{\rho}$, and so V is some sort of geometric object that joins together all the formal deformation theory of the various $\bar{\rho}$. The next question, then, is exactly what sort of geometric object is V ? While the V_Q are finitely presented affine schemes, their colimit V is more infinitary, and we need to study what sort of geometry we can get when we take this sort of colimit.

So what does V look like? It could involve taking a countable disjoint union of varieties: this is still geometric, and it's locally of finite presentation with countably many components (recall each of the V_Q are finitely presented). A slightly more complicated possibility is taking a not-necessarily disjoint union of varieties, each of which is an irreducible (though not necessarily connected) component of their union. This isn't of finite presentation, but it is still a scheme, and it's still quite geometric.

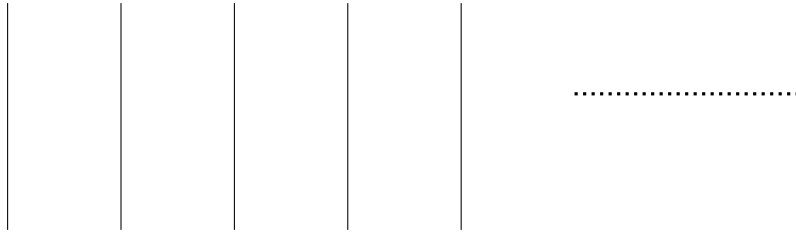


FIGURE 1. An infinite union of connected components

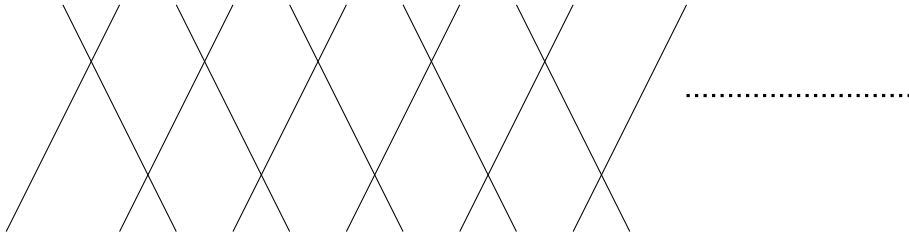


FIGURE 2. An infinite union of irreducible components

But it could fail to be a scheme! For instance, take $\mathbf{A}_{\mathbf{F}_p}^1$ and glue a copy of $\mathbf{A}_{\mathbf{F}_p}^1$ to each closed point.

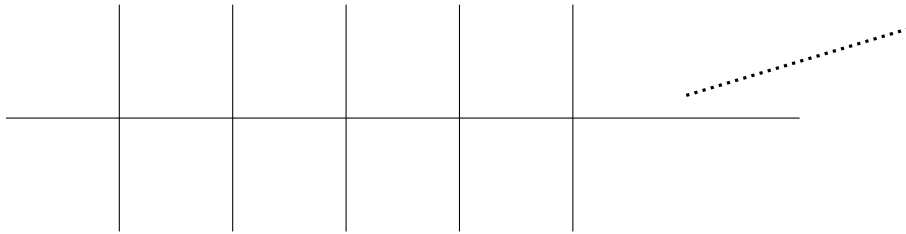


FIGURE 3. An Ind-scheme which is not a scheme

We refer the reader to [EG21, §4.3] for many more examples of Ind-schemes that aren't schemes.

So if V is of one of the first two types, and is still a scheme, then we'd be in good shape because we can do geometry, but otherwise we haven't done much to simplify our original problem. In fact, it really depends on the characteristic of our coefficient field: in characteristic p we will see that things behave badly, while in characteristic 0, we will see that we get a well-behaved formal scheme.

1.3. Analysis of V when $d = 1$ and $K = \mathbf{Q}_p$. For simplicity, assume $p > 2$. In this case, we use local class field theory to replace $G_{\mathbf{Q}_p}$ with

$$G_{\mathbf{Q}_p}^{\text{ab}} = \widehat{\mathbf{Q}_p^\times} = \mathbf{Z}_p^\times \times p^{\widehat{\mathbf{Z}}}.$$

Taking the Weil–Deligne subgroup gives

$$\text{WD}_{\mathbf{Q}_p}^{\text{ab}} = \mathbf{Q}_p^\times = \mathbf{Z}_p^\times \times p^{\mathbf{Z}}.$$

In this case, the pro- p wild ramification part is a copy of \mathbf{Z}_p , embedded in $\text{WD}_{\mathbf{Q}_p}^{\text{ab}}$ via $1 + p\mathbf{Z}_p \hookrightarrow \mathbf{Z}_p^\times$. So we choose a sequence $Q = Q_n$ so that

$$\text{WD}_{\mathbf{Q}_p}^{\text{ab}}/Q_n = (\mathbf{Z}/p^{n+1})^\times \times p^{\mathbf{Z}} \cong C_{p^n} \times C_{p-1} \times p^{\mathbf{Z}}.$$

Thus

$$V_{Q_n} = \text{Spec } \mathbf{Z}[x]/(x^{p^n} - 1) \times \text{Spec } \mathbf{Z}[x]/(x^{p-1} - 1) \times \mathbf{G}_{m, \mathbf{Z}}.$$

- (1) To see what happens away from p , we can invert p on this scheme and see what we get:

$$\begin{aligned} V_{\mathbf{Z}[1/p]} &= \varinjlim_n V_{Q_n, \mathbf{Z}[1/p]} \\ &= (\varinjlim_n \text{Spec } \mathbf{Z}[1/p][x]/(x^{p^n} - 1)) \times \text{Spec } \mathbf{Z}[1/p][x]/(x^{p-1} - 1) \times \mathbf{G}_{m, \mathbf{Z}[1/p]}. \end{aligned}$$

If we now base change to $k = \overline{\mathbf{Q}}$ or $k = \overline{\mathbf{F}}_\ell$ for $\ell \neq p$, then $(x^{p^n} - 1)$ is separable, so we are just adding more and more closed points as we go further along the directed system, and so we basically get infinitely many copies of $\text{Spec } \mathbf{Z}[x]/(x^{p-1} - 1) \times \mathbf{G}_m$, indexed by p th roots of unity in k .

- (2) But instead if we base change to \mathbf{Z}_p , then the closed immersions $V_{Q_n, \mathbf{Z}_p} \hookrightarrow V_{Q_{n+1}, \mathbf{Z}_p}$ are actually nilpotent thickenings, and in fact we end up with

$$V_{\mathbf{Z}_p} = \text{Spf } \mathbf{Z}_p[[T]] \times \text{Spec } \mathbf{Z}[x]/(x^{p-1} - 1) \times \mathbf{G}_m.$$

Now $\text{Spf } \mathbf{Z}_p[[T]]$ is a nice Noetherian formal scheme; be warned that this is specific to dimension 1. Indeed already in dimension 2 the situation becomes much more complicated, as we will see shortly.

Before turning to the case of dimensions 2 and higher, we give some more examples of Ind-schemes. First, here's a formal scheme that isn't Noetherian: Take $\mathbf{A}_{\overline{\mathbf{F}}_p}^1$ and add a extra formal direction at each closed point on the line (one can do this finitely many times and then take a filtered colimit over the finite steps). This is an affine formal scheme $\text{Spf } A$ for A some topological ring, but A is not Noetherian. More precisely, we let $\overline{\mathbf{F}}_p = \{a_0, a_1, \dots\}$ and set $A := \varprojlim_j A_j$ where

$$A_j = \overline{\mathbf{F}}_p[x] \times_{\overline{\mathbf{F}}_p} \overline{\mathbf{F}}_p[[x - a_0]] \times_{\overline{\mathbf{F}}_p} \cdots \times_{\overline{\mathbf{F}}_p} \overline{\mathbf{F}}_p[[x - a_j]].$$

Now here’s something even worse. Consider the subsheaf of $\mathbf{A}_{\overline{\mathbf{F}}_p}^2$ given by taking a “horizontal” $\mathbf{A}_{\overline{\mathbf{F}}_p}^1$ and adding a “vertical” $\mathbf{A}_{\overline{\mathbf{F}}_p}^1$ through each closed point, as in Figure 3, and then infinitesimally thickening each of the vertical lines (in $\mathbf{A}_{\overline{\mathbf{F}}_p}^2$) in the horizontal direction, as in Figure 4.



FIGURE 4. An Ind-scheme with thickened vertical lines

Then this is not even a formal scheme, and it’s actually hard to tell this apart from \mathbf{A}^2 , because it has the same closed points and versal rings.

We could even take the above construction and then delete the point at the origin. Call this V_* ; we will return to it shortly.

1.4. Analysis of V in general. For general d and general K , $V_Q/\mathbf{Z}[1/p]$ is reduced, local complete intersection (so Cohen–Macaulay), and flat/ $\mathbf{Z}[1/p]$ of relative dimension d^2 [DHKM20, Zhu20], and

$$V_Q/\mathbf{Z}[1/p] \hookrightarrow V_{Q'}/\mathbf{Z}[1/p]$$

is just adding connected components. Note in particular that the local deformation rings of mod ℓ residual representations are just computed using this variety by looking at the complete local rings at the stalks of the corresponding points.

Now let’s switch back to the characteristic p setting. Let $K = \mathbf{Q}_p$, $d = 2$, and assume that $p > 2$. Let’s just study the behaviour of V in characteristic p ; let’s also fix the determinant of our two-dimensional Galois representations to be ω^i , where ω denotes the mod p cyclotomic character, and we choose $1 \leq i \leq p - 2$. We write $V_{\overline{\mathbf{F}}_p}^{\det=\omega^i}$ for the resulting Ind-scheme.

Let’s begin by considering the $\overline{\mathbf{F}}_p$ -valued points of $V_{\overline{\mathbf{F}}_p}^{\det=\omega^i}$; these correspond to $\overline{\rho} : \mathrm{WD}_{\mathbf{Q}_p} \rightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ of determinant ω^i . There is one irreducible such $\overline{\rho}$ up to isomorphism, namely $\mathrm{Ind}_{\mathbf{Q}_{p^2}}^{\mathbf{Q}_p} \omega_2^i$. Because of the choice of framing, this will not correspond to a literal single closed point of $V_{\overline{\mathbf{F}}_p}^{\det=\omega^i}$, but rather to an SL_2 orbit.² This orbit *will* be (the underlying reduced subscheme of) a connected component of $V_{\overline{\mathbf{F}}_p}^{\det=\omega^i}$.

Any reducible $\overline{\rho}$ will be an extension of mod p characters of $\mathrm{WD}_{\mathbf{Q}_p}$, and any such character is an unramified twist of a power of ω . Any family of such extensions will induce a corresponding family of semi-simplifications (more precisely, of associated pseudo-characters), and the powers of the cyclotomic character that appear will be locally constant, and thus constant on a given connected component of the family. We will restrict attention to those $\overline{\rho}$ for which $\overline{\rho}^{\mathrm{ss}} = \mathrm{ur}_{\lambda^{-1}\omega^i} \oplus \mathrm{ur}_{\lambda}$, for some $\lambda \in \overline{\mathbf{F}}_p^\times$,

²“Fixing the determinant” means choosing an isomorphism of $\mathrm{WD}_{\mathbf{Q}_p}$ -representations $\wedge^2 \rho \xrightarrow{\sim} \omega^i$, and so the “change of frame” group is reduced from GL_2 to SL_2 .

and in fact to those $\bar{\rho}$ which are an extension of ur_λ by $\text{ur}_{\lambda^{-1}\omega^i}$ (rather than an extension in the opposite order). Any such non-split extension is unique up to isomorphism (equivalently, the corresponding Ext^1 is one-dimensional) unless $i = 1$ and $\lambda = \pm 1$; in this latter case there is a one-dimensional space of *peu ramifiée* extensions, and it is the $\bar{\rho}$ classified by this space that we will consider.

There is actually a p -adic Hodge-theoretic framework that describes exactly these $\bar{\rho}$ that we are singling out (the irreducible $\bar{\rho}$, and the particular reducible $\bar{\rho}$ that we have described); namely, they are *Fontaine–Laffaille* with Hodge–Tate weights $\{0, i\}$ (and with determinant ω^i). And what we will now describe is the Ind-scheme $V_{/\mathbf{F}_p}^{\text{FL}, \det=\omega^i}$ which classifies families of $\text{WD}_{\mathbf{Q}_p}$ -representations over finite type \mathbf{F}_p -algebras A that are two-dimensional, have determinant fixed to be ω^i , and which are Fontaine–Laffaille with Hodge weights $\{0, i\}$ when pushed forward to any Artinian quotient of A .

Actually, to avoid grappling with the details of framings, we will describe the quotient stack³ $[V_{/\mathbf{F}_p}^{\text{FL}, \det=\omega^i}/\text{SL}_2]$. To state the answer, recall the subsheaf V_\star of $\mathbf{A}_{/\mathbf{F}_p}^2$ described above. There is an action of \mathbf{G}_m on this substack obtained by restricting the following action of \mathbf{G}_m on \mathbf{A}^2 :

$$a \cdot (x, y) = (x, a^2 y).$$

1.4.1. Proposition. *There is an isomorphism of stacks*

$$[V_{/\mathbf{F}_p}^{\text{FL}, \det=\omega^i}/\text{SL}_2] \xrightarrow{\sim} [V_\star/\mathbf{G}_m].$$

Before sketching the proof, we note that V_\star is rather nasty. It is Zariski dense in $\mathbf{A}^2 \setminus \{0\}$ — indeed, it has the same set of $\overline{\mathbf{F}_p}$ -points, and the same versal rings at these points, as $\mathbf{A}^2 \setminus \{0\}$ — but it is disconnected, and it looks more like a formal scheme than an honest scheme like $\mathbf{A}^2 \setminus \{0\}$. Given this, it makes sense to ask: can one study some related but different moduli problem so as to actually get $\mathbf{A}^2 \setminus \{0\}$ (or, rather, the stack $[(\mathbf{A}^2 \setminus \{0\})/\mathbf{G}_m]$) as the answer? Well, you can! If you study families of Fontaine–Laffaille *modules* (the “linear algebra perspective” in p -adic Hodge theory) instead of *representations* then this is exactly what you get. The inclusion

$$[V_{/\mathbf{F}_p}^{\text{FL}, \det=\omega^i}/\text{SL}_2] \xrightarrow{\sim} [V_\star/\mathbf{G}_m] \hookrightarrow [(\mathbf{A}^2 \setminus \{0\})/\mathbf{G}_m]$$

identifies the source with the locus in the Fontaine–Laffaille moduli stack over which the Fontaine–Laffaille functor from Fontaine–Laffaille modules to Galois representations (or, more generally, to $\text{WD}_{\mathbf{Q}_p}$ -representations) can be applied so as to obtain genuine $\text{WD}_{\mathbf{Q}_p}$ -representations.

Sketch of proof of Prop. 1.4.1. As already noted, both the source and the target of the claimed isomorphism are disconnected: the source is the disjoint union of the reducible and irreducible loci, while the target is the disjoint union of $[(V_\star \cap (\mathbf{G}_m \times \mathbf{A}^1))/\mathbf{G}_m]$ and its complement (which is a single closed point $[(0 \times \mathbf{G}_m)/\mathbf{G}_m] = [\bullet/\mu_2]$ with a μ_2 -inertia group). We check the isomorphism on each connected component separately. On the irreducible locus, it amounts to the fact that the mod p Fontaine–Laffaille deformation ring of $\text{Ind}_{\mathbf{Q}_p}^{\mathbf{Q}_p} \omega_2^i$ is isomorphic to $\mathbf{F}_p[[x]]$. On the reducible locus, there is more to check, but the key point is that if $\bar{\rho}$ is a non-split extension of ur_λ by $\text{ur}_{\lambda^{-1}\omega^i}$, for some $\lambda \in \overline{\mathbf{F}_p}^\times$, which is trivial on some

³See the next lecture for more about stacks!

given open subgroup Q of $P_{\mathbf{Q}_p}$, then the degree of λ over \mathbf{F}_p is bounded in terms of the index $[P_{\mathbf{Q}_p} : Q]$. \square

This example of $V^{\text{FL}, \det=\omega^i}$ is illustrative of the overall behaviour of V : when working over \mathbf{F}_p or \mathbf{Z}_p , it is nasty as soon as $d > 1$. In the special case of $V^{\text{FL}, \det=\omega^i}$, the moduli space of Fontaine–Laffaille modules served as a much more pleasant improvement. Fontaine–Laffaille theory doesn’t apply to general mod p or p -adic Galois representations (even for \mathbf{Q}_p , but even more so if K/\mathbf{Q}_p is ramified). But Fontaine’s theory of (φ, Γ) -modules does! And in fact the moduli space V does sit inside a moduli space of (φ, Γ) -modules, and as we will see, this latter space is a lot nicer. Thus we are motivated to define moduli stacks of (φ, Γ) -modules and do a thorough study of their geometry.

LECTURE 2. STACKS, ETC.

In this lecture we give a very brief introduction to the idea of a formal algebraic stack; details can be found in [Eme].

2.1. Functors of points. There’s a tension between rings and topological spaces in algebraic geometry, going back to Diophantus and Descartes. Grothendieck’s formulation encompasses both, by giving you a topological space, but also giving you a sheaf of rings. You then solve equations by studying morphisms of schemes.

To study algebraic spaces and stacks, it is helpful to take the functorial point of view, which encompasses the theory of schemes, but gives you a framework to extend it. This perspective starts with the following observation.

2.1.1. Lemma (Yoneda, Grothendieck). *A scheme X can be thought of as a (pre)sheaf $\text{Sch} \xrightarrow{X} \text{Set}$ by taking $Y \mapsto \text{Hom}(Y, X)$. In fact it’s a Zariski/étale/fppf sheaf.*

But we can consider more general sheaves. For instance consider a system $X_1 \hookrightarrow X_2 \hookrightarrow \dots$ with closed immersions as transition maps, and define a sheaf

$$X := \varinjlim X_i$$

This is an Ind-scheme. Often, this will not actually be a scheme: e.g. embed a point in a line in a plane, in 3-space, etc. This gives you some “infinite dimensional affine space”, which is not a scheme: note the identity map from this space to itself doesn’t factor through one of the finite steps, whereas a map from a quasi-compact scheme does always factor in this way.

This last remark illustrates one way that finiteness/quasi-compactness assumptions are used in our theory; they help us to detect finite-dimensional parts of possibly infinite-dimensional objects.

2.2. Algebraic Spaces. Another example of a more general sheaf is an algebraic space. If X is an fppf sheaf, and U is a scheme and $U \rightarrow X$ is a morphism, then we say that the morphism $U \rightarrow X$ is “representable by schemes” if in any fibre product diagram

$$\begin{array}{ccc} T \times_X U & \longrightarrow & U \\ \downarrow & & \downarrow \\ T & \longrightarrow & X \end{array}$$

for which T is a scheme, then $T \times_X U$ is also a scheme. It usually suffices to check this for a certain subclass of T (e.g. affine schemes). You could also require that the morphism $U \rightarrow X$ is surjective or étale, which are both properties that one can define via base change to schemes.

2.2.1. Definition. An *algebraic space* is an *fppf* sheaf X which admits a map $U \rightarrow X$ whose source is a scheme and which is representable by schemes, surjective, and étale.

There is an obvious notion of an Ind-algebraic space.

2.3. Formal algebraic spaces. We want to be able to talk about formal algebraic spaces, and formal algebraic stacks.

Suppose A is a complete topological ring with a countable basis of open neighborhoods of 0 consisting of ideals I_n . Assume that every $A/I_{n+1} \rightarrow A/I_n$ is a nilpotent thickening. Then we define

$$\mathrm{Spf} A = \varinjlim \mathrm{Spec} A/I_n.$$

For instance if $A = \mathbf{Z}_p$ and $I_n = p^n$, then we get $\mathrm{Spf} \mathbf{Z}_p$. This is a proper subsheaf of $\mathrm{Spec} \mathbf{Z}_p$. Interestingly, properties like flatness, reducedness, Cohen–Macaulayness need the ring A , and can't be seen on the quotients A/I_n . For instance, \mathbf{Z}_p is reduced, while each $\mathbf{Z}/p^n\mathbf{Z}$ for $n > 1$ is not.

2.3.1. Definition. A *formal algebraic space* is an *fppf* sheaf X which admits a map $\bigsqcup_i U_i \rightarrow X$ for which each U_i is an affine formal scheme, and such that the map is representable by algebraic spaces, étale and surjective.

Then if $X = \varinjlim X_i$, where the X_i are algebraic spaces and the maps are nilpotent thickenings, one can show that X is a formal algebraic space. Vice versa, you can write a formal algebraic space using an Ind-construction.

2.4. Stacks. Stacks are “sheaves” of groupoids. Some (small) groupoids are contractible, and are equivalent to their underlying sets, but some are not. For example, consider the category with one object x and two morphisms $\{1_x, \sigma\}$ where $\sigma^2 = 1_x$: this is not contractible. The moral is that in sets, equality is a *property*, but in higher category theory and the theory of stacks, equality is exhibited by an isomorphism which is some *extra data*.

Really we want to say that a stack is a “2-sheaf”, which means a 2-functor from the category of schemes to the 2-category of groupoids, satisfying a 2-categorical analogue of the sheaf condition. This is technically possible to do, but is a bit complicated once you start trying to work out all the coherence conditions you need, and usually involves making some kind of choices of pullbacks.

On the other hand, one way to avoid this is to use the formalism of *categories fibered in groupoids*, which [Sta] does, and which we will do. For example, a morphism of stacks is a fully faithful embedding of categories fibered in groupoids. An occasionally annoying terminological issue is that “isomorphism of stacks” means an equivalence of categories fibered in groupoids. From now on we will typically denote stacks by calligraphic letters \mathcal{X} , and schemes (or more generally sheaves) by roman letters X .

There are some interesting things you get from the “extra data” involved in having to *choose* an isomorphism. For example, the diagonal morphism $\Delta : \mathcal{X} \rightarrow \mathcal{X} \times \mathcal{X}$ is a monomorphism if and only if \mathcal{X} is equivalent to something that lands in

sets (in other words, if and only if \mathcal{X} can be regarded as being a usual sheaf of sets). In general, the diagonal really tells you something about isomorphisms between objects in the groupoids. For instance, if T is a scheme and $f : T \rightarrow \mathcal{X} \times \mathcal{X}$ is a morphism, induced by the pair of morphisms $f_1, f_2 : T \rightrightarrows \mathcal{X}$, then the pull-back of Δ along f will be the Isom sheaf classifying isomorphisms between the elements f_1 and f_2 of the groupoid $\mathcal{X}(T)$.

2.5. Algebraic and Formal Algebraic Stacks. We make the following definitions.

2.5.1. Definition. We say that an *fppf* stack \mathcal{X} is an *algebraic stack* if there exists a morphism $U \rightarrow \mathcal{X}$ whose source is a scheme, and which is representable by algebraic spaces, surjective, and smooth. Note that in this case asking for the morphism to be étale is strictly stronger than asking for smooth, and defines the notion of a Deligne–Mumford stack, which we will not make use of.

2.5.2. Example. Recall the schemes V_Q from the previous lecture, whose A -valued points correspond to representations $\mathrm{WD}_K/Q \rightarrow \mathrm{GL}_d(A)$. Usually in Galois representation theory, we want to study representations up to isomorphism, which means that we should quotient out by the conjugation action of GL_d . This motivates the consideration of the quotient stack $[V_Q/\mathrm{GL}_d]$, which is an algebraic stack, and is the *moduli stack* of d -dimensional WD_K/Q -representations.

We can also form $[V/\mathrm{GL}_d] = \varinjlim_Q [V_Q/\mathrm{GL}_d]$, which is an Ind-algebraic stack. If we work over $\mathbf{Z}[1/p]$, then we saw in the last lecture that V becomes a usual scheme (though not of finite type), and so then $[V/\mathrm{GL}_d]$ is again an algebraic stack.

If we work over \mathbf{Z}_p , then we've seen that V , and so also $[V/\mathrm{GL}_d]$, is not a very good object. As we've already said, we will replace this moduli stack of WD_K -representations by a moduli stack of étale (φ, Γ) -modules. This won't be an algebraic stack, but it will be a formal algebraic stack, in the following sense.

2.5.3. Definition. We say that an *fppf* stack \mathcal{X} is a *formal algebraic stack* if there exists a morphism $\bigsqcup_i U_i \rightarrow \mathcal{X}$ which is representable by algebraic spaces, surjective, and smooth, with the U_i being affine formal schemes.

If \mathcal{X} is qcqs, then we may write $\mathcal{X} = \varinjlim \mathcal{X}_i$, where \mathcal{X}_i are algebraic stacks and the transition maps are thickenings. The converse is true if \varinjlim is countably indexed and the transition maps are thickenings.

If $f : \mathcal{X} \rightarrow \mathcal{Y}$ is a morphism of stacks that is representable by algebraic stacks (in the obvious sense), you can ascribe geometric properties to f . In particular, if $\Delta : \mathcal{Y} \rightarrow \mathcal{Y} \times \mathcal{Y}$ is representable by algebraic spaces, and \mathcal{X} is an algebraic stack, then any morphism $\mathcal{X} \rightarrow \mathcal{Y}$ is representable by algebraic stacks.

Now imagine \mathcal{X} is an algebraic stack, locally of finite type over an excellent locally Noetherian scheme S , and consider a map $\mathcal{X} \rightarrow \mathcal{F}$, where \mathcal{F} is an *fppf* stack for which $\Delta_{\mathcal{F}}$ is representable by algebraic spaces. Then as remarked in the previous paragraph, the morphism $\mathcal{X} \rightarrow \mathcal{F}$ is representable by algebraic stacks. In particular, it makes sense to ask that $\mathcal{X} \rightarrow \mathcal{F}$ be proper. In practice, \mathcal{X} will be a stack of Breuil–Kisin modules of bounded height, and \mathcal{F} will be a stack of étale φ -modules.

We would like to form the “scheme-theoretic image” of \mathcal{F} in \mathcal{X} . Morally, this amounts to contracting the proper equivalence relation $\mathcal{F} \times_{\mathcal{X}} \mathcal{F}$ on \mathcal{F} . In general

you wouldn't expect contracting a proper equivalence relation to give you something algebraic. However, we have the following theorem, the main result of [EG21].

2.5.4. Theorem (Scheme Theoretic Images). *Assume given an fppf sheaf \mathcal{F} which is limit preserving and whose diagonal is representable by algebraic spaces. Suppose that \mathcal{X} is an algebraic stack, and that $\mathcal{X} \rightarrow \mathcal{F}$ is a proper morphism. Assume further that \mathcal{F} admits versal rings at all finite type points, and that these rings satisfy the following effectivity property with respect to the image of \mathcal{X} : if $\mathrm{Spf} R \rightarrow \mathcal{F}$ is a morphism realizing the pro-Artinian local ring R as a versal ring to \mathcal{F} , if we pull \mathcal{X} back over this morphism to obtain a proper map $\mathcal{X}_R \rightarrow \mathrm{Spf} R$, and if we let $\mathrm{Spf} S \hookrightarrow \mathrm{Spf} R$ be the scheme-theoretic image of this map, then the composite morphism $\mathrm{Spf} S \rightarrow \mathrm{Spf} R \rightarrow \mathcal{F}$ factors through $\mathrm{Spec} S$.*

Then there exists an algebraic closed substack $\mathcal{Z} \hookrightarrow \mathcal{F}$ such that $\mathcal{X} \rightarrow \mathcal{F}$ factors through a morphism $\mathcal{X} \rightarrow \mathcal{Z}$ which is proper and scheme theoretically dominant. Furthermore, the closed substack \mathcal{Z} is uniquely determined by these properties, and we refer to it as the scheme-theoretic image of the morphism $\mathcal{X} \rightarrow \mathcal{F}$.

We will use this theorem repeatedly in our context to construct the Ind-algebraic stacks that we want.

LECTURE 3. DEFINITIONS RELATED TO ÉTALE φ -MODULES

In this lecture we give the definitions of the various kinds of φ -modules with coefficients that we make use of; in the following lecture, we will begin to use these definitions to define our moduli stacks, and to prove their basic properties.

3.1. Rings. We begin with some material from [EG19, §2.1]. We fix a finite extension K of \mathbf{Q}_p with ring of integers \mathcal{O}_K and residue field k , and regard K as a subfield of some fixed algebraic closure $\overline{\mathbf{Q}_p}$ of \mathbf{Q}_p . We let \mathbf{C} denote the completion of $\overline{\mathbf{Q}_p}$. It is a perfectoid field, whose tilt \mathbf{C}^\flat is a complete non-archimedean valued perfect field of characteristic p . If F is a perfectoid closed subfield of \mathbf{C} , then its tilt F^\flat is a closed, and perfect, subfield of \mathbf{C}^\flat . We will only need to consider the following two examples of F (other than \mathbf{C} itself), which arise from the theories of (φ, Γ) -modules and Breuil–Kisin modules.

3.1.1. Example (The cyclotomic case). We write $K(\zeta_{p^\infty})$ to denote the extension of K obtained by adjoining all p -power roots of unity. It is an infinite degree Galois extension of K , whose Galois group is naturally identified with an open subgroup of \mathbf{Z}_p^\times . We let K_{cyc} denote the unique subextension of $K(\zeta_{p^\infty})$ whose Galois group over K is isomorphic to \mathbf{Z}_p (so K_{cyc} is the “cyclotomic \mathbf{Z}_p -extension” of K). If we let $\widehat{K}_{\mathrm{cyc}}$ denote the closure of K_{cyc} in \mathbf{C} , then $\widehat{K}_{\mathrm{cyc}}$ is a perfectoid subfield of \mathbf{C} .

3.1.2. Example (The Kummer case). If we choose a uniformizer π of K , as well as a compatible system of p -power roots π^{1/p^n} of π (here, “compatible” has the obvious meaning, namely that $(\pi^{1/p^{n+1}})^p = \pi^{1/p^n}$), then we define $K_\infty = K(\pi^{1/p^\infty}) := \bigcup_n K(\pi^{1/p^n})$. If we let \widehat{K}_∞ denote the closure of K_∞ in \mathbf{C} , then \widehat{K}_∞ is again a perfectoid subfield of \mathbf{C} .

3.1.3. Remark. Let $L = K_{\mathrm{cyc}}$ or K_∞ . Then $\widehat{L} \otimes_L \overline{\mathbf{Q}_p}$ is an algebraic closure of \widehat{L} , so that the absolute Galois groups of L and \widehat{L} are canonically identified. The action of G_L on $\overline{\mathbf{Q}_p}$ extends to an action on \mathbf{C} , and by a theorem of Ax–Tate–Sen [Ax70], we have $\mathbf{C}^{G_L} = \widehat{L}$.

If \mathcal{O}_F denotes the ring of integers in F (with $F = \mathbf{C}$ or one of the two possibilities just discussed), then \mathcal{O}_F^b is the ring of integers in F^b . Since F^b and \mathcal{O}_F^b are perfect, we may form their rings of (truncated) Witt vectors $W(\mathcal{O}_F^b)$, $W(F^b)$, $W_a(F^b)$, $W_a(\mathcal{O}_F^b)$ (where $a \geq 1$ is an integer). We write $\mathbf{A}_{\text{inf}} := W(\mathcal{O}_{\mathbf{C}}^b)$.

We always consider these rings of Witt vectors as topological rings with the so-called *weak topology*, which admits the following description: if x is any element of \mathcal{O}_F^b of positive valuation, and if $[x]$ denotes the Teichmüller lift of x , then we endow $W_a(\mathcal{O}_F^b)$ with the $[x]$ -adic topology, so that $W(\mathcal{O}_F^b)$ is then endowed with the $(p, [x])$ -adic topology. The topology on $W_a(F^b)$ is then characterized by the fact that $W_a(\mathcal{O}_F^b)$ is an open subring (and the topology on $W(F^b)$ is the inverse limit topology).

While we could formulate all of our results in terms of φ -modules over these rings of Witt vectors, for the purposes of proving our structural results about the stacks (and for connecting to the usual theories of (φ, Γ) -modules and Breuil–Kisin modules), we need to consider various smaller (in particular Noetherian) subrings. These come from the Fontaine–Wintenberger theory of the field of norms, but we will skip over this and go straight to the definitions.

Firstly, in the Kummer case, we set $\mathfrak{S} = W(k)[[u]]$, with a φ -semi-linear endomorphism φ determined by $\varphi(u) = u^p$, and let $\mathcal{O}_{\mathcal{E}}$ be the p -adic completion of $\mathfrak{S}[1/u]$. The choice of compatible system of p -power roots of π gives an element $\pi^{1/p^\infty} \in \mathcal{O}_{\widehat{K_\infty}}^b$, and there is a continuous φ -equivariant embedding

$$\mathfrak{S} \hookrightarrow W(\mathcal{O}_{\widehat{K_\infty}}^b)$$

sending $u \mapsto [\pi^{1/p^\infty}]$. This embedding extends to a continuous φ -equivariant embedding

$$\mathcal{O}_{\mathcal{E}} \hookrightarrow W(\widehat{(\widehat{K_\infty})}^b).$$

Now consider the cyclotomic case, where we write $\widetilde{\Gamma}_K := \text{Gal}(K(\zeta_{p^\infty})/K)$, and the cyclotomic character induces an embedding $\chi : \widetilde{\Gamma}_K \hookrightarrow \mathbf{Z}_p^\times$. Consequently, there is an isomorphism $\widetilde{\Gamma}_K \cong \Gamma_K \times \Delta$, where $\Gamma_K \cong \mathbf{Z}_p$ and Δ is finite. We have $K_{\text{cyc}} = (K(\zeta_{p^\infty}))^\Delta$. We will usually write Γ for Γ_K from now on.

We now make a simplifying assumption for the purposes of exposition: assume that K/\mathbf{Q}_p is unramified. We will keep this assumption in place for the rest of this lecture and the following two lectures. (All of our results hold for general K/\mathbf{Q}_p , but there is a subtlety: for general K/\mathbf{Q}_p , there is no φ -stable analogue of the ring \mathbf{A}_K^+ . This means that in many arguments in [EG19] we reduce to the unramified case by somewhat technical although essentially straightforward arguments. We want to avoid these complications in these lectures.)

If we choose a compatible system of p^n th roots of 1, then these give rise in the usual way to an element $\varepsilon \in \widehat{K(\zeta_{p^\infty})}^b$. There is then a continuous embedding

$$W(k)[[T]] \hookrightarrow W(\mathcal{O}_{\widehat{K(\zeta_{p^\infty})}}^b)$$

(the source being endowed with its (p, T) -adic topology, and the target with its weak topology), defined via $T \mapsto [\varepsilon] - 1$. We denote the image of this embedding by $(\mathbf{A}'_K)^+$. This embedding extends to an embedding

$$W(\widehat{k}(\widehat{T})) \hookrightarrow W(\widehat{(\widehat{K(\zeta_{p^\infty})}}^b)$$

(here the source is the p -adic completion of the Laurent series ring $W(k)((T))$), whose image we denote by \mathbf{A}'_K . Write $T'_K \in \mathbf{A}'_K$ for the image of T . The actions of φ and $\gamma \in \tilde{\Gamma}_K$ on $T'_K \in \mathbf{A}'_K$ are given by the explicit formulae

$$(3.1.4) \quad \varphi(T'_K) = (1 + T'_K)^p - 1,$$

$$(3.1.5) \quad \gamma(1 + T'_K) = (1 + T'_K)^{\chi(\gamma)},$$

where $\chi : \tilde{\Gamma}_K \rightarrow \mathbf{Z}_p^\times$ again denotes the cyclotomic character. We set $\mathbf{A}_K := (\mathbf{A}'_K)^\Delta$, $T_K = \text{tr}_{\mathbf{A}'_K/\mathbf{A}_K}(T'_K)$, and $\mathbf{A}_K^+ = W(k)[[T_K]]$; then we have $\mathbf{A}_K = W(k)((T_K))^\wedge$, and \mathbf{A}_K^+ is (φ, Γ_K) -stable. We have $\varphi(T_K) \in T_K \mathbf{A}_K^+$, and $g(T_K) \in T_K \mathbf{A}_K^+$ for all $g \in \Gamma_K$. From now on we will often write T for T_K .

3.1.6. *Remark.* We could equally well work with the rings \mathbf{A}'_K and the theory of $(\varphi, \tilde{\Gamma})$ -modules, but it is often convenient to work with the procyclic group Γ and not have to carry around the finite group Δ .

3.1.7. *Remark.* Note that in both the Kummer and cyclotomic settings, we are considering rings which are abstractly a power series ring over the Witt vectors, equipped with a lift of Frobenius. This will mean that various foundational parts of the theory can be developed in parallel for the two cases.

3.2. **Coefficients.** We now consider coefficients. Recall that if A is a p -adically complete \mathbf{Z}_p -algebra, then A is said to be *topologically of finite type over \mathbf{Z}_p* if it can be written as a quotient of a restricted formal power series ring in finitely many variables $\mathbf{Z}_p\langle\langle X_1, \dots, X_n \rangle\rangle$; equivalently, if and only if A/p is a finite type \mathbf{F}_p -algebra ([FK18, §0, Prop. 8.4.2]). In particular, if A is a \mathbf{Z}/p^a -algebra for some $a \geq 1$, then A is topologically of finite type over \mathbf{Z}_p if and only if it is of finite type over \mathbf{Z}/p^a . Our coefficient rings will always be assumed to be topologically of finite type over \mathbf{Z}_p , and they will usually also be (finite type) \mathbf{Z}/p^a -algebras for some $a \geq 1$.

3.2.1. *Remark.* Our stacks are all limit preserving, so that values on A -valued points (for any A) are determined by their values on A which are (topologically) of finite type. It therefore does not cause us any problems to restrict to coefficients A of this kind. In addition many of our arguments with φ -modules require this assumption on A , and we do not know if the results hold without it.

Our various coefficient rings are all defined by taking completed tensor products. More precisely, if $a \geq 1$, let v denote an element of the maximal ideal of $W_a(\mathcal{O}_{\mathbf{C}}^b)$ whose image in $\mathcal{O}_{\mathbf{C}}^b$ is non-zero. We then set

$$W_a(\mathcal{O}_{\mathbf{C}}^b)_A = W_a(\mathcal{O}_{\mathbf{C}}^b) \widehat{\otimes}_{\mathbf{Z}_p} A := \varprojlim_i (W_a(\mathcal{O}_{\mathbf{C}}^b) \otimes_{\mathbf{Z}_p} A) / v^i$$

(so that the indicated completion is the v -adic completion). Note that any two choices of v induce the same topology on $W_a(\mathcal{O}_{\mathbf{C}}^b) \otimes_{\mathbf{Z}_p} A$, so that $W_a(\mathcal{O}_{\mathbf{C}}^b) \widehat{\otimes}_{\mathbf{Z}_p} A$ is well-defined independent of the choice of v . We then define

$$W_a(\mathbf{C}^b)_A = W_a(\mathbf{C}^b) \widehat{\otimes}_{\mathbf{Z}_p} A := W_a(\mathcal{O}_{\mathbf{C}}^b)_A[1/v];$$

this ring is again well-defined independently of the choice of v . We set

$$W(\mathcal{O}_{\mathbf{C}}^b)_A = W(\mathcal{O}_{\mathbf{C}}^b) \widehat{\otimes}_{\mathbf{Z}_p} A := \varprojlim_a W_a(\mathcal{O}_{\mathbf{C}}^b)_A,$$

and similarly

$$W(\mathbf{C}^b)_A = W(\mathbf{C}^b) \widehat{\otimes}_{\mathbf{Z}_p} A := \varprojlim_a W_a(\mathbf{C}^b)_A.$$

In keeping with our notation above, we will usually write

$$\mathbf{A}_{\text{inf},A} := W(\mathcal{O}_{\mathbf{C}}^b)_A.$$

We define the imperfect coefficient rings in the Kummer and cyclotomic cases in the same way, but we can be more explicit: we have $\mathfrak{S}_A = (W(k) \otimes_{\mathbf{Z}_p} A)[[u]]$, and $\mathcal{O}_{\mathcal{E},A}$ is the p -adic completion of $\mathfrak{S}_A[1/u]$, and similarly for $\mathbf{A}_{K,A}^+$ and $\mathbf{A}_{K,A}$.

3.2.2. A digression on flatness and completion. Since \mathbf{A}_{inf} is not Noetherian, we have to be a bit careful when taking completions, and for example it does not seem to be obvious that the natural map $\mathfrak{S}_A \rightarrow \mathbf{A}_{\text{inf},A}$ is injective. The following is [EG19, Prop. 2.2.11], and is proved using [FGK11, §5] and [BMS18, Rem. 4.31].

3.2.3. Proposition. *Suppose that $A \rightarrow B$ is a flat homomorphism of finite type $\mathbf{Z}/p^a\mathbf{Z}$ -algebras for some $a \geq 1$. Then all the maps in the following diagram are flat. Furthermore the vertical arrows are all injections, while the horizontal arrows are all faithfully flat (and so in particular also injections). If $A \rightarrow B$ is furthermore faithfully flat, then the same is true of the diagonal arrows.*

$$\begin{array}{ccccccc}
& & \mathbf{A}_{K,B}^+ & \longrightarrow & \mathbf{A}_{\text{inf},B} & \longleftarrow & \mathfrak{S}_B \\
& \nearrow & \downarrow & & \nearrow & & \downarrow \\
\mathbf{A}_{K,A}^+ & \longrightarrow & \mathbf{A}_{\text{inf},A} & \longleftarrow & \mathfrak{S}_A & \longrightarrow & \mathfrak{S}_B \\
& \downarrow & \downarrow & & \downarrow & & \downarrow \\
& & \mathbf{A}_{K,B} & \longrightarrow & W(\mathbf{C}^b)_B & \longleftarrow & \mathcal{O}_{\mathcal{E},B} \\
& \nearrow & \downarrow & & \nearrow & & \downarrow \\
\mathbf{A}_{K,A} & \longrightarrow & W(\mathbf{C}^b)_A & \longleftarrow & \mathcal{O}_{\mathcal{E},A} & \longrightarrow & \mathcal{O}_{\mathcal{E},B}
\end{array}$$

It is straightforward to show that φ extends to a continuous endomorphism of $\mathbf{A}_{\text{inf},A}$, $W(\mathcal{O}_{\mathbf{C}}^b)_A$, and so on; and the natural action of G_K on these rings is continuous, as is the action of Γ on $\mathbf{A}_{K,A}^+$ and $\mathbf{A}_{K,A}$.

3.3. φ -modules, Breuil–Kisin(–Fargues) modules, (φ, Γ) -modules, and Galois representations. Let R be a \mathbf{Z}_p -algebra, equipped with a ring endomorphism φ , which is congruent to the (p -power) Frobenius modulo p . If M is an R -module, we write

$$\varphi^* M := R \otimes_{R, \varphi} M.$$

3.3.1. Definition. An *étale φ -module over R* is a finite R -module M , equipped with a φ -semi-linear endomorphism $\varphi_M : M \rightarrow M$, which has the property that the induced R -linear morphism

$$\Phi_M : \varphi^* M \xrightarrow{1 \otimes \varphi_M} M$$

is an isomorphism. A morphism of étale φ -modules is a morphism of the underlying R -modules which commutes with the morphisms Φ_M . We say that M is *projective*

(resp. *free*) if it is projective of constant rank (resp. free of constant rank) as an R -module.

We apply this in particular with $R = \mathbf{A}_{K,A}$ or $\mathcal{O}_{\mathcal{E},A}$. In the former case, an *étale* (φ, Γ) -module is an étale φ -module over $\mathbf{A}_{K,A}$, equipped with a commuting continuous semilinear action of Γ . In both cases there is a relationship with Galois representations as follows; the case without coefficients is due to Fontaine (and was the motivation for introducing (φ, Γ) -modules in the first place), and the version with coefficients is due to Dee [Dee01].

Let $\widehat{\mathbf{A}}_K^{\text{ur}}$ denote the p -adic completion of the ring of integers of the maximal unramified extension of $\mathbf{A}_K[1/p]$ in $W(\mathbf{C}^b)[1/p]$; this is preserved by the natural actions of φ and G_K on $W(\mathbf{C}^b)[1/p]$. Define $\widehat{\mathbf{A}}_{K,A}^{\text{ur}}$ as usual. Then if A is an Artinian \mathbf{Z}_p -algebra, we have an equivalence of categories between the category of finite projective étale (φ, Γ) -modules M with A -coefficients, and the category of finite free A -modules V with a continuous action of G_K , given by the functors

$$\begin{aligned} V &\mapsto (\widehat{\mathbf{A}}_{K,A}^{\text{ur}} \otimes_A V)^{G_{K_{\text{cyc}}}}, \\ M &\mapsto (\widehat{\mathbf{A}}_{K,A}^{\text{ur}} \otimes_{\mathbf{A}_{K,A}} M)^{\varphi=1}. \end{aligned}$$

Taking limits (and forming $\widehat{\mathbf{A}}_{K,A}^{\text{ur}}$ with respect to the \mathfrak{m}_A -adic topology), we can extend this to the case that A is complete local Noetherian, or $A = \overline{\mathbf{F}}_p$ or $\overline{\mathbf{Z}}_p$; this is important later on for identifying the versal rings of our moduli stacks with Galois deformation rings. Note though that for more general A (e.g. $A = \mathbf{F}_p[X]$), it is not the case that there is an equivalence between (φ, Γ) -modules and Galois representations (*cf.* the discussion in Lecture 1.4).

There is also a version of this in the Kummer setting: if $\mathcal{O}_{\widehat{\mathcal{E}}^{\text{ur}}}$ is the p -adic completion of the ring of integers in the maximal unramified extension of $\text{Frac}(\mathcal{O}_{\mathcal{E}})$ in $W(\mathbf{C}^b)[1/p]$, then for A as above, we have an equivalence of categories between étale φ -modules over $\mathcal{O}_{\mathcal{E},A}$ and G_{K_∞} -representations, given by

$$\begin{aligned} V &\mapsto (\widehat{\mathcal{O}}_{\widehat{\mathcal{E}}^{\text{ur}},A} \otimes_A V)^{G_{K_\infty}}, \\ M &\mapsto (\widehat{\mathcal{O}}_{\widehat{\mathcal{E}}^{\text{ur}},A} \otimes_{\mathcal{O}_{\mathcal{E},A}} M)^{\varphi=1}. \end{aligned}$$

In the next few lectures we will define moduli stacks of étale φ -modules and étale (φ, Γ) -modules. One of the key tools in proving the basic properties of these stacks is to study the corresponding stacks of *finite height* φ -modules over $\mathbf{A}_{K,A}^+$ or \mathfrak{S}_A . We'll do this more generally in the next lecture, but for now we'll just have the following definition, which is important in the crystalline and semistable theory. (Actually, because K/\mathbf{Q}_p is assumed unramified, the finite height modules over $\mathbf{A}_{K,A}^+$ have an intrinsic utility for describing crystalline representations in terms of *Wach modules*. However, this doesn't extend to general K/\mathbf{Q}_p , or to the semistable case, so we will not make any use of this property of Wach modules in these lectures.)

3.3.2. Definition. Let $E(u)$ be the minimal polynomial over $W(k)$ of our fixed uniformiser π of K . We define a (projective) Breuil–Kisin module (resp. a Breuil–Kisin–Fargues module) of height at most h with A -coefficients to be a finitely generated projective \mathfrak{S}_A -module (resp. $\mathbf{A}_{\text{inf},A}$ -module) \mathfrak{M} , equipped with a φ -semilinear morphism $\varphi : \mathfrak{M} \rightarrow \mathfrak{M}$, with the property that the corresponding morphism $\Phi_{\mathfrak{M}} : \varphi^* \mathfrak{M} \rightarrow \mathfrak{M}$ is injective, with cokernel killed by $E(u)^h$. If \mathfrak{M} is a Breuil–Kisin module then $\mathbf{A}_{\text{inf},A} \otimes_{\mathfrak{S}_A} \mathfrak{M}$ is a Breuil–Kisin–Fargues module.

Note that our Breuil–Kisin(–Fargues) modules are “effective”, i.e. φ -stable. This means that we will end up only considering non-negative Hodge–Tate weights; the general case follows by twisting by a sufficiently large power of $E(u)$, so this is harmless.

The connection between Breuil–Kisin(–Fargues) modules and crystalline/semistable Galois representations will be explained in Lecture 6.

3.4. Almost Galois descent. Let F be a closed perfectoid subfield of \mathbf{C} . The following is [EG19, Thm. 2.4.1], which is proved by a somewhat elaborate argument making use of almost Galois descent.

3.4.1. Theorem. *Let A be a finite type \mathbf{Z}/p^a -algebra, for some $a \geq 1$. The inclusion $W(F^b)_A \rightarrow W(\mathbf{C}^b)_A$ is a faithfully flat morphism of Noetherian rings, and the functor $M \mapsto W(\mathbf{C}^b)_A \otimes_{W(F^b)_A} M$ induces an equivalence between the category of finitely generated projective $W(F^b)_A$ -modules and the category of finitely generated projective $W(\mathbf{C}^b)_A$ -modules endowed with a continuous semi-linear G_F -action. A quasi-inverse functor is given by $N \mapsto N^{G_F}$.*

We will find this result very useful later on, because it lets us combine the Kummer and cyclotomic settings. The key extra ingredient (which also takes some work) is that when we have the additional structure of a φ -module, we can descend from the perfect coefficient ring $W(F^b)_A$ (with $F = \widehat{K}_{\text{cyc}}$ or $F = \widehat{K}_\infty$) to the imperfect coefficient rings $\mathbf{A}_{K,A}$ and $\mathcal{O}_{\mathcal{E},A}$ respectively, in the following way.

3.4.2. Definition. Let A be a p -adically complete \mathbf{Z}_p -algebra. An étale (G_K, φ) -module with A -coefficients (resp. an étale $(G_{K_{\text{cyc}}}, \varphi)$ -module with A -coefficients, resp. an étale (G_{K_∞}, φ) -module with A -coefficients) is by definition a finitely generated $W(\mathbf{C}^b)_A$ -module M equipped with an isomorphism of $W(\mathbf{C}^b)_A$ -modules

$$\Phi_M : \varphi^* M \xrightarrow{\sim} M,$$

and a $W(\mathbf{C}^b)_A$ -semi-linear action of G_K (resp. $G_{K_{\text{cyc}}}$, resp. G_{K_∞}), which is continuous and commutes with Φ_M . We say that M is projective if it is projective of constant rank as a $W(\mathbf{C}^b)_A$ -module.

The following is [EG19, Prop. 2.7.4].

3.4.3. Proposition. *Let A be a finite type \mathbf{Z}/p^a -algebra for some $a \geq 1$.*

- (1) *The functor $M \mapsto W(\mathbf{C}^b)_A \otimes_{\mathbf{A}_{K,A}} M$ is an equivalence between the category of finite projective étale φ -modules over $\mathbf{A}_{K,A}$ and the category of finite projective étale $(G_{K_{\text{cyc}}}, \varphi)$ -modules with A -coefficients.*

It induces an equivalence of categories between the category of finite projective étale (φ, Γ_K) -modules with A -coefficients and the category of finite projective étale (G_K, φ) -modules with A -coefficients.

- (2) *The functor $M \mapsto W(\mathbf{C}^b)_A \otimes_{\mathcal{O}_{\mathcal{E},A}} M$ is an equivalence of categories between the category of finite projective étale φ -modules over $\mathcal{O}_{\mathcal{E},A}$ and the category of finite projective étale (G_{K_∞}, φ) -modules with A -coefficients.*

LECTURE 4. MODULI STACKS OF φ -MODULES

We now define our first moduli stacks, using the objects introduced in the previous lecture. We maintain our assumption (made purely for the purposes of exposition) that K/\mathbf{Q}_p is unramified.

4.1. Results of Pappas–Rapoport. We begin our discussion of stacks with φ -modules; the more complicated case of (φ, Γ) -modules will be built on this. These stacks were first studied by Pappas–Rapoport [PR09] in the context of Breuil–Kisin modules; they showed (among other things) that there are algebraic stacks of Breuil–Kisin modules, and that the morphisms to the stacks of étale φ -modules are well-behaved. These results were built upon in [EG21, §5], which showed that the stack of étale φ -modules is itself well-behaved, and in particular is Ind-algebraic.

We begin by recalling the results of Pappas–Rapoport (in a slightly more general context, where φ is not necessarily given by $\varphi(u) = u^p$; this makes some arguments messier but doesn't change any of the key points).

4.1.1. Situation. Fix a finite extension k/\mathbf{F}_p and write $\mathbf{A}^+ := W(k)[[T]]$. Write \mathbf{A} for the p -adic completion of $\mathbf{A}^+[1/T]$.

If A is a p -adically complete \mathbf{Z}_p -algebra, we write $\mathbf{A}_A^+ := (W(k) \otimes_{\mathbf{Z}_p} A)[[T]]$; we equip \mathbf{A}_A^+ with its (p, T) -adic topology, so that it is a topological A -algebra (where A has the p -adic topology). Let \mathbf{A}_A be the p -adic completion of $\mathbf{A}_A^+[1/T]$, which we regard as a topological A -algebra by declaring \mathbf{A}_A^+ to be an open subalgebra.

Let φ be a ring endomorphism of \mathbf{A} which is congruent to the $(p$ -power) Frobenius endomorphism modulo p , and satisfies $\varphi(\mathbf{A}^+) \subseteq \mathbf{A}^+$.

By [EG21, Lem. 5.2.2 and 5.2.5] and [EG19, Lem. 3.2.5, 3.2.6], φ is faithfully flat, and induces the usual Frobenius on $W(k)$, and it extends uniquely to an A -linear continuous endomorphism of \mathbf{A}_A^+ and \mathbf{A}_A .

Note in particular that the formation of \mathfrak{S}_A , $\mathcal{O}_{\mathcal{E}, A}$, $\mathbf{A}_{K, A}^+$ and $\mathbf{A}_{K, A}$ above are particular instances of this construction.

Fix a polynomial $F \in W(k)[T]$ which is congruent to a positive power of T modulo p . (For example, if we are working in the Breuil–Kisin setting, we will take $T = u$, $\varphi(u) = u^p$, and $F = E(u)$.)

In order to know that the various categories in groupoids that we will define are actually stacks, we use the following result of Drinfeld [Dri06, Thm. 3.11] (see [EG21, Thm. 5.1.18] for the precise statements given here).

4.1.2. Theorem. *The following notions are local for the fpqc topology on $\mathrm{Spec} A$.*

- (1) *A finitely generated projective \mathbf{A}_A -module.*
- (2) *A projective \mathbf{A}_A -module of rank d .*
- (3) *A finitely generated projective \mathbf{A}_A -module which is fpqc locally free of rank d .*
- (4) *A finitely generated projective \mathbf{A}_A^+ -module.*
- (5) *A projective \mathbf{A}_A^+ -module of rank d .*
- (6) *A finitely generated projective \mathbf{A}_A^+ -module which is fpqc locally free of rank d .*

4.1.3. Remark. More precisely, saying that the notion of a finitely generated projective \mathbf{A}_A -module is local for the fpqc topology on $\mathrm{Spec} A$ means the following (and the meanings of the other statements in Theorem 4.1.2 are entirely analogous):

If A' is any faithfully flat A -algebra, set $A'' := A' \otimes_A A'$. Then the category of finitely generated projective \mathbf{A}_A -modules is canonically equivalent to the category of finitely generated projective $\mathbf{A}_{A'}$ -modules M' which are equipped with an isomorphism

$$M' \otimes_{\mathbf{A}_{A'}, a \mapsto 1 \otimes a} \mathbf{A}_{A''} \xrightarrow{\sim} M' \otimes_{\mathbf{A}_{A'}, a \mapsto a \otimes 1} \mathbf{A}_{A''}$$

which satisfies the usual cocycle condition.

If we fix integers $a, d \geq 1$, then by Theorem 4.1.2 we may define an *fpgc* stack in groupoids \mathcal{R}_d^a over $\mathrm{Spec} \mathbf{Z}/p^a$ as follows: For any \mathbf{Z}/p^a -algebra A , we define $\mathcal{R}_d^a(A)$ to be the groupoid of étale φ -modules M over \mathbf{A}_A which are projective of rank d . There is a closely related version of this considered in [PR09], namely $\mathcal{R}_{d,\mathrm{free}}^a$, where we demand that M is furthermore *fpgc*-locally free. (In fact, we don't know whether or not this is a consequence of M being projective, so we don't know whether $\mathcal{R}_{d,\mathrm{free}}^a = \mathcal{R}_d^a$.) In either case, if $A \rightarrow B$ is a morphism of \mathbf{Z}_p -algebras, and M is an object of $\mathcal{R}_d^a(A)$, then the pull-back of M to $\mathcal{R}_d^a(B)$ is defined to be the tensor product $\mathbf{A}_B \otimes_{\mathbf{A}_A} M$.

From now on we regard \mathcal{R}_d^a as an *fppf* stack over \mathbf{Z}/p^a . By [Sta, Tag 04WV], we may also regard the stack \mathcal{R}_d^a as an *fppf* stack over \mathbf{Z}_p , and as a varies, we may form the 2-colimit $\mathcal{R} := \varinjlim_a \mathcal{R}_d^a$, which is again an *fppf* stack over \mathbf{Z}_p .

The following definition generalises that of a Breuil–Kisin module.

4.1.4. Definition. Let h be a non-negative integer. A φ -module of *F-height at most h over \mathbf{A}_A^+* is a pair $(\mathfrak{M}, \varphi_M)$ consisting of a finitely generated T -torsion free \mathbf{A}_A^+ -module \mathfrak{M} , and a φ -semi-linear map $\varphi_{\mathfrak{M}} : \mathfrak{M} \rightarrow \mathfrak{M}$, with the further properties that if we write

$$\Phi_{\mathfrak{M}} := 1 \otimes \varphi_{\mathfrak{M}} : \varphi^* \mathfrak{M} \rightarrow \mathfrak{M},$$

then $\Phi_{\mathfrak{M}}$ is injective, and the cokernel of $\Phi_{\mathfrak{M}}$ is killed by F^h .

A φ -module of *finite F-height over \mathbf{A}_A^+* is a φ -module of *F-height at most h* for some $h \geq 0$. A morphism of φ -modules is a morphism of the underlying \mathbf{A}_A^+ -modules which commutes with the morphisms $\Phi_{\mathfrak{M}}$.

We say that a φ -module of finite *F-height* is projective of rank d if it is a finitely generated projective \mathbf{A}_A^+ -module of constant rank d .

If we fix integers $a, d \geq 1$ and an integer $h \geq 0$, then we may define an *fpgc* stack in groupoids $\mathcal{C}_{d,h}^a$ over $\mathrm{Spec} \mathbf{Z}/p^a$ as follows: For any \mathbf{Z}/p^a -algebra A , we define $\mathcal{C}_{d,h}^a(A)$ to be the groupoid of φ -modules of *F-height at most h over \mathbf{A}_A^+* which are projective of rank d .

Just as for the stack \mathcal{R}_d^a , we may and do also regard the stack $\mathcal{C}_{d,h}^a$ as an *fppf* stack over \mathbf{Z}_p , and we then, allowing a to vary, define $\mathcal{C}_{d,h} := \varinjlim_a \mathcal{C}_{d,h}^a$, obtaining an *fppf* stack over $\mathrm{Spf} \mathbf{Z}_p$. There are canonical morphisms $\mathcal{C}_{d,h}^a \rightarrow \mathcal{R}_d^a$ and $\mathcal{C}_{d,h} \rightarrow \mathcal{R}_d$ given by tensoring with \mathbf{A}_A over \mathbf{A}_A^+ . One can show that any projective \mathbf{A}_A^+ -module is even Zariski locally free, so these morphisms factor through the Pappas–Rapoport versions $\mathcal{R}_{d,\mathrm{free}}^a$ (resp. $\mathcal{R}_{d,\mathrm{free}}$). Reassuringly, we have the following lemma [EG19, Lem. 3.1.3].

4.1.5. Lemma. *If A is a p -adically complete \mathbf{Z}_p -algebra, then there is a canonical equivalence between the groupoid of morphisms $\mathrm{Spf} A \rightarrow \mathcal{R}_d$ and the groupoid of rank d étale φ -modules over \mathbf{A}_A ; and there is a canonical equivalence between the groupoid of morphisms $\mathrm{Spf} A \rightarrow \mathcal{C}_{d,h}$ and the groupoid of φ -modules of rank d and *F-height at most h over \mathbf{A}_A^+ .**

The following is essentially [PR09, Thm. 2.1 (a), Cor. 2.6].

4.1.6. Theorem.

- (1) *The stack $\mathcal{C}_{d,h}^a$ is an algebraic stack of finite presentation over $\mathrm{Spec} \mathbf{Z}/p^a$, with affine diagonal.*

- (2) The morphism $\mathcal{C}_{d,h}^a \rightarrow \mathcal{R}_{d,\text{free}}^a$ is representable by algebraic spaces, proper, and of finite presentation.
- (3) The diagonal morphism $\Delta : \mathcal{R}_{d,\text{free}}^a \rightarrow \mathcal{R}_{d,\text{free}}^a \times_{\mathbf{Z}/p^a} \mathcal{R}_{d,\text{free}}^a$ is representable by algebraic spaces, affine, and of finite presentation.

We very briefly indicate the main ideas of the proof. One of the key points is the following [PR09, Prop. 2.2] (see [EG21, Lem. 5.2.9] for this version). Assume that A is a $\mathbf{Z}/p^a\mathbf{Z}$ -algebra. For $n \geq 0$, write

$$U_n = 1 + T^n M_d(\mathbf{A}_A^+),$$

$$V_n = \{A \in \text{GL}_d(\mathbf{A}_A) \mid A, A^{-1} \in T^{-n} M_d(\mathbf{A}_A^+)\}.$$

4.1.7. Lemma. *For each $m \geq 0$ there is an $n(m) \geq 0$ (implicitly depending also on a) such that if $n \geq n(m)$, then:*

- (1) For each $g \in U_n$, $B \in V_m$, there is a unique $h \in U_n$ such that $g^{-1}B\varphi(g) = h^{-1}B$.
- (2) For each $h \in U_n$, $B \in V_m$ there is a unique $g \in U_n$ such that $g^{-1}B\varphi(g) = h^{-1}B$.

In each case the uniqueness statement is easy, and the existence is proved by a (slightly tricky) successive approximation argument. Given the lemma, the first part of Theorem 4.1.6 follows from standard results about the affine Grassmannian: the point is that working locally, the data of a finite height \mathbf{A}_A^+ -module is the data of the matrix in $M_d(\mathbf{A}_A^+)$ of the corresponding linearized φ , and the finite height condition implies that this is in V_n for some n depending on a and h . Changes of basis correspond to φ -conjugacy, and Lemma 4.1.7 lets us replace φ -conjugacy by left multiplication provided we work in a sufficiently deep congruence subgroup, and this is enough to get the result.

The rest of Theorem 4.1.6 is also proved by reducing to explicit statements about matrices over \mathbf{A}_A . The key point in each case is to obtain a bound on the T -adic poles in some matrix entries, and the bound on the height h gives a bound on how different the denominators of g and $\varphi(g)$ can be (for some matrix g), which can be played off against the fact that the poles of $\varphi(g)$ have approximately p times the order of the poles of g .

We then deduce that Theorem 4.1.6 holds with $\mathcal{R}_{d,\text{free}}^a$ replaced by \mathcal{R}_d^a ; roughly speaking the idea is to regard a projective \mathbf{A}_A -module as a pair consisting of an *fpqc*-locally free projective \mathbf{A}_A -module of higher rank, together with an idempotent. We also make use of the following technical statement [EG21, Lem. 5.1.23]: if M is a finitely generated projective $A((u))$ -module, then there exists an $n_0 \geq 1$ such that for all $n \geq n_0$, M is Nisnevich (and in particular *fpqc*) locally free as an $A((u^n))$ -module. In particular we find it useful to regard a projective $A((u))$ -module as a locally free $A((u^n))$ -module together with an action of u .

4.2. Scheme-theoretic images. The main result of [EG21] is the following [EG21, Thm. 5.4.20].

4.2.1. Theorem. *\mathcal{R}_d^a is a limit preserving Ind-algebraic stack, whose diagonal is representable by algebraic spaces, affine, and of finite presentation.*

The key point is of course the Ind-algebraicity. In fact we show that $\mathcal{R}_d^a = \varinjlim_h \mathcal{R}_{d,h}^a$, where $\mathcal{R}_{d,h}^a$ is the scheme-theoretic image (in the sense of Theorem 2.5.4) of the morphism $\mathcal{C}_{d,h}^a \rightarrow \mathcal{R}_d^a$. It is at least intuitively reasonable that we have

$\mathcal{R}_d^a = \varinjlim_h \mathcal{R}_{d,h}^a$ (this says that étale locally, every étale φ -module comes from a finite height φ -module, which in the free case is immediate by just choosing a basis and scaling it by a sufficiently large power of F), and the hard part is to show that $\mathcal{R}_{d,h}^a$ is actually an algebraic stack.

The key point in applying Theorem 2.5.4 is to show that $\mathcal{R}_{d,h}^a$ admits *effective* Noetherian versal rings. Apart from the effectivity, this is at least morally quite straightforward: the versal rings for \mathcal{R}_d^a correspond to (infinite-dimensional) unrestricted Galois deformation rings, and those for $\mathcal{R}_{d,h}^a$ are closely related to their finite height analogues, which are Noetherian. In particular, if we are in the Kummer setting and we take $h = 1$, then \mathcal{R}_d^a really admits the unrestricted deformation rings for G_{K_∞} as versal rings, and $\mathcal{R}_{d,1}^a$ admits a version of the height 1 deformation rings, which can be identified with finite flat deformation rings for G_K (which are in particular Noetherian).

The real work is in proving the effectivity, i.e. that the versal morphism $\mathrm{Spf} R \rightarrow \mathcal{R}_{d,h}^a$ can be upgraded to a morphism $\mathrm{Spec} R \rightarrow \mathcal{R}_{d,h}^a$. The key point here is to check that the composite morphism $\mathrm{Spf} R \rightarrow \mathcal{R}_d^a$ comes from a morphism $\mathrm{Spec} R \rightarrow \mathcal{R}_d^a$; it is then relatively straightforward to check that this morphism factors through $\mathcal{R}_{d,h}^a$, using arguments which are very similar to those that we will give in Section 5.3. Slightly more explicitly, having a morphism $\mathrm{Spf} R \rightarrow \mathcal{R}_d^a$ means that we have an étale φ -module over the \mathfrak{m}_R -adic completion of \mathbf{A}_R (i.e. where we are allowed to have infinite Laurent tails, as long as they tend to zero \mathfrak{m}_R -adically), and we need to show that this arises from a projective étale φ -module over \mathbf{A}_R itself.

The reason this is true is that there is a universal (not necessarily projective) φ -module \mathfrak{M}_R over \mathbf{A}_R^\dagger of F -height at most h , arising as the pushforward of the universal φ -module over $\mathcal{C}_{d,h}^a$. We then set $M_{R^h} = \mathfrak{M}_{R^h}[1/T]$, which gives us a (not necessarily projective) étale φ -module over \mathbf{A}_R , whose \mathfrak{m}_R -adic completion agrees with the φ -module obtained from $\mathrm{Spf} R \rightarrow \mathcal{R}_d^a$. (This compatibility with the \mathfrak{m}_R -adic completion is not completely obvious, but ultimately follows from the theorem on formal functions.) It remains to check that M_R is projective; this is immediate from the following general theorem ([EG21, Thm. 5.5.20]).

4.2.2. Theorem. *Let R be a complete Noetherian local \mathbf{Z}/p^a -algebra with maximal ideal \mathfrak{m} , let M be an étale φ -module over \mathbf{A}_R , and suppose that the \mathfrak{m} -adic completion \widehat{M} is projective, or equivalently, free (over the \mathfrak{m} -adic completion of \mathbf{A}_R). Then M itself is projective (over \mathbf{A}_R).*

Finally, passing to the limit over $a \geq 1$ and using some results of [Eme], we obtain the following.

4.2.3. Corollary.

- (1) $\mathcal{C}_{d,h}$ is a p -adic formal algebraic stack of finite presentation over $\mathrm{Spf} \mathbf{Z}_p$, with affine diagonal.
- (2) \mathcal{R}_d is a limit preserving Ind-algebraic stack, whose diagonal is representable by algebraic spaces, affine, and of finite presentation.
- (3) The morphism $\mathcal{C}_{d,h} \rightarrow \mathcal{R}_d$ is representable by algebraic spaces, proper, and of finite presentation.
- (4) The diagonal morphism $\Delta : \mathcal{R}_d \rightarrow \mathcal{R}_d \times_{\mathrm{Spf} \mathbf{Z}_p} \mathcal{R}_d$ is representable by algebraic spaces, affine, and of finite presentation.

LECTURE 5. STACKS OF (φ, Γ) -MODULES

5.1. **Moduli stacks of (φ, Γ) -modules.** Recall that for simplicity we are always assuming that K/\mathbf{Q}_p is unramified when we discuss (φ, Γ) -modules.

5.1.1. **Definition.** We let $\mathcal{X}_{K,d}$ denote the moduli stack of projective étale (φ, Γ) -modules of rank d . More precisely, if A is a p -adically complete \mathbf{Z}_p -algebra, then we define $\mathcal{X}_{K,d}(A)$ (i.e., the groupoid of morphisms $\mathrm{Spf} A \rightarrow \mathcal{X}_{K,d}$) to be the groupoid of projective étale (φ, Γ) -modules of rank d with A -coefficients, with morphisms given by isomorphisms. If $A \rightarrow B$ is a morphism of complete \mathbf{Z}_p -algebras, and M is an object of $\mathcal{X}_{K,d}(A)$, then the pull-back of M to $\mathcal{X}_{K,d}(B)$ is defined to be the tensor product $\mathbf{A}_{K,B} \otimes_{\mathbf{A}_{K,A}} M$. It again follows from Theorem 4.1.2 that $\mathcal{X}_{K,d}$ is indeed a stack.

Note that by the equivalence between (φ, Γ) -modules and G_K -representations over finite \mathbf{Z}_p -algebras, the $\overline{\mathbf{F}}_p$ -points of $\mathcal{X}_{K,d}$ are in bijection with representations $\overline{\rho} : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{F}}_p)$. More generally, fix a point $\mathrm{Spec} \mathbf{F} \rightarrow \mathcal{X}_{K,d}$ for some finite field \mathbf{F} , giving rise to a continuous representation $\overline{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})$, and let $R_{\overline{\rho}}^{\square}$ denote the universal framed deformation $W(\mathbf{F})$ -algebra for lifts of $\overline{\rho}$. Then it is easy to check that the natural morphism $\mathrm{Spf} R_{\overline{\rho}}^{\square} \rightarrow \mathcal{X}_{K,d}$ (which again is defined via the equivalence between (φ, Γ) -modules and G_K -representations over finite \mathbf{Z}_p -algebras) is versal.

One of the main results that we will prove is that $\mathcal{X}_{K,d}$ is a Noetherian formal algebraic stack. However, the proof of this is quite involved, and in this lecture we establish the preliminary result that $\mathcal{X}_{K,d}$ is an Ind-algebraic stack, which we deduce from the results that we have proved for stacks of étale φ -modules.

5.1.2. **Definition.** We let $\mathcal{R}_{K,d}$ (frequently abbreviated to \mathcal{R}_d) denote the moduli stack of rank d projective étale φ -modules, taking \mathbf{A} to be \mathbf{A}_K .

We begin by studying Γ -actions on our φ -modules. We choose a topological generator γ of Γ , and let $\Gamma_{\mathrm{disc}} := \langle \gamma \rangle$; so $\Gamma_{\mathrm{disc}} \cong \mathbf{Z}$. Note that since Γ_{disc} is dense in Γ , in order to endow M with the structure of an étale (φ, Γ) -module, it suffices to equip M with a continuous action of Γ_{disc} (where we equip Γ_{disc} with the topology induced on it by Γ).

There is a canonical action of Γ_{disc} on \mathcal{R}_d (that is, a canonical morphism $\gamma : \mathcal{R}_d \rightarrow \mathcal{R}_d$): if M is an object of $\mathcal{R}_d(A)$, then $\gamma(M)$ is given by $\gamma^*M := \mathbf{A}_{K,A} \otimes_{\gamma, \mathbf{A}_{K,A}} M$. Then we set

$$\mathcal{R}_d^{\Gamma_{\mathrm{disc}}} := \mathcal{R}_d \times_{\Delta, \mathcal{R}_d \times \mathcal{R}_d, \Gamma_{\gamma}} \mathcal{R}_d,$$

where Δ is the diagonal of \mathcal{R}_d and Γ_{γ} is the graph of γ , so that $\Gamma_{\gamma}(x) = (x, \gamma(x))$.

We claim that $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}}$ is nothing other than the moduli stack of projective étale φ -modules of rank d equipped with a semi-linear action of Γ_{disc} . This is an exercise in unwinding the usual construction of the 2-fibre product: $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}}$ consists of tuples (x, y, α, β) , with x, y being objects of \mathcal{R}_d , and $\alpha : x \xrightarrow{\sim} y$ and $\beta : \gamma(x) \xrightarrow{\sim} y$ being isomorphisms. This is equivalent to the category fibred in groupoids given by pairs (x, ι) consisting of an object x of \mathcal{R}_d and an isomorphism $\iota : \gamma(x) \xrightarrow{\sim} x$. Thus an object of $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}}(A)$ is a projective étale φ -module of rank d with A -coefficients M , together with an isomorphism of φ -modules $\iota : \gamma^*M \xrightarrow{\sim} M$; and this isomorphism is precisely the data of a semi-linear action of $\Gamma_{\mathrm{disc}} = \langle \gamma \rangle$ on M , as required.

Since \mathcal{R}_d is an Ind-algebraic stack, so is $\mathcal{R}_d^{\Gamma_{\text{disc}}}$; indeed it follows from Corollary 4.2.3 that $\mathcal{R}_d^{\Gamma_{\text{disc}}}$ is a limit preserving Ind-algebraic stack, whose diagonal is representable by algebraic spaces, affine, and of finite presentation.

Restricting the Γ -action on an étale (φ, Γ) -module to Γ_{disc} , we obtain a morphism $\mathcal{X}_{K,d} \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}}$, which is fully faithful. Thus $\mathcal{X}_{K,d}$ may be regarded as a substack of $\mathcal{R}_d^{\Gamma_{\text{disc}}}$; in particular, we deduce that its diagonal is representable by algebraic spaces, affine, and of finite presentation. Although $\mathcal{X}_{K,d}$ is a substack of the Ind-algebraic stack $\mathcal{R}_d^{\Gamma_{\text{disc}}}$, it is not a closed substack, but should rather be thought of as a certain formal completion (as in the discussion of the rank one case in Lecture 1.3); in particular, since it is not a closed substack, we cannot immediately deduce the Ind-algebraicity of $\mathcal{X}_{K,d}$ from the Ind-algebraicity of $\mathcal{R}_d^{\Gamma_{\text{disc}}}$. Instead, we will argue as in the proof that \mathcal{R}_d is Ind-algebraic, and exhibit $\mathcal{X}_{K,d}$ as the scheme-theoretic image of an Ind-algebraic stack.

From now on we will typically drop K from the notation, simply writing \mathcal{X}_d , \mathcal{R}_d and so on. To go further we need to understand the continuity condition in the definition of \mathcal{X}_d more carefully. It is easy to check that we have $\gamma(T) - T \in (p, T)T\mathbf{A}_A^+$, and using this one can check the following special case of [EG19, Lem. D.24]:

5.1.3. Lemma. *Suppose that A is a \mathbf{Z}/p^a -algebra for some $a \geq 1$. Let M be a finite projective \mathbf{A}_A -module, equipped with a semi-linear action of Γ_{disc} . Then the following are equivalent:*

- (1) *The action of Γ_{disc} extends to a continuous action of Γ .*
- (2) *For any lattice $\mathfrak{M} \subseteq M$, there exists $s \geq 0$ such that $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$. (Here a lattice is a finitely generated $\mathbf{A}_{K,A}^+$ -submodule $\mathfrak{M} \subseteq M$ whose $\mathbf{A}_{K,A}$ -span is M .)*
- (3) *For some lattice $\mathfrak{M} \subseteq M$ and some $s \geq 0$, we have $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$.*
- (4) *The action of $\gamma - 1$ on $M \otimes_{\mathbf{Z}/p^a} \mathbf{F}_p$ is topologically nilpotent.*

It is easy to use this criterion, together with the fact that $\mathcal{R}_d^{\Gamma_{\text{disc}}}$ is limit preserving, to show that \mathcal{X}_d is limit preserving; see [EG19, Lem. 3.2.19].

5.2. Weak Wach modules. In this section we introduce the notion of a weak Wach module of height at most h and level at most s . These will play a purely technical auxiliary role for us, and will be used only in order to show that \mathcal{X}_d is an Ind-algebraic stack; we won't use their relation to crystalline representations.

By Lemma 5.1.3, if A is a \mathbf{Z}/p^a -algebra for some $a \geq 1$, and \mathfrak{M} is a rank d projective φ -module of T -height $\leq h$ over A , such that $\mathfrak{M}[1/T]$ is equipped with a semi-linear action of Γ_{disc} , then this action extends to a continuous action of Γ if and only if for some $s \geq 0$ we have $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$. This motivates the following definition.

5.2.1. Definition. A rank d projective weak Wach module of T -height $\leq h$ and level $\leq s$ is a rank d projective φ -module \mathfrak{M} over $\mathbf{A}_{K,A}^+$, which is of T -height $\leq h$, such that $\mathfrak{M}[1/T]$ is equipped with a semilinear action of Γ_{disc} which satisfies $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$.

5.2.2. Definition. We let $\mathcal{W}_{d,h}$ denote the moduli stack of rank d projective weak Wach modules of T -height $\leq h$. We let $\mathcal{W}_{d,h,s}$ denote the substack of rank d projective weak Wach modules of T -height $\leq h$ and level $\leq s$.

We will next show that the stacks $\mathcal{W}_{d,h,s}$ are p -adic formal algebraic stacks of finite presentation over $\mathrm{Spf} \mathbf{Z}_p$. Since the canonical morphism $\varinjlim_s \mathcal{W}_{d,h,s} \rightarrow \mathcal{W}_{d,h}$ is an isomorphism (by definition), this will show in particular that $\mathcal{W}_{d,h}$ is an Ind-algebraic stack; we will also see that the transition maps in this inductive limit are closed immersions.

Recall that we have the p -adic formal algebraic stack $\mathcal{C}_{d,h}$ classifying rank d projective φ -modules over $\mathbf{A}_{K,A}^+$ of T -height at most h . We consider the fibre product $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h}$, where the map $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \rightarrow \mathcal{R}_d$ is the canonical morphism given by forgetting the Γ_{disc} action; this is the moduli stack of rank d projective φ -modules \mathfrak{M} over $\mathbf{A}_{K,A}^+$ of T -height at most h , equipped with a semi-linear action of Γ_{disc} on $\mathfrak{M}[1/T]$. It follows from Corollary 4.2.3 that $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h}$ is a p -adic formal algebraic stack of finite presentation over $\mathrm{Spf} \mathbf{Z}_p$.

Restricting the Γ -action on a weak Wach module to Γ_{disc} , we may regard $\mathcal{W}_{d,h}$ as a substack of $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h}$. The following is [EG19, Prop. 3.3.5].

5.2.3. Proposition. *For $s \geq 1$, the morphism*

$$\mathcal{W}_{d,h,s} \longrightarrow \mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h}$$

is a closed immersion of finite presentation. In particular, each of the stacks $\mathcal{W}_{d,h,s}$ is a p -adic formal algebraic stack of finite presentation over $\mathrm{Spf} \mathbf{Z}_p$; and for each $s' \geq s$, the canonical monomorphism $\mathcal{W}_{d,h,s} \hookrightarrow \mathcal{W}_{d,h,s'}$ is a closed immersion of finite presentation.

The proof of this is fairly straightforward: by definition, we need to show that the condition that $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$ is a closed condition, and is determined by finitely many equations. We do this by reducing to the free case and considering the equations on the level of matrices.

5.3. \mathcal{X}_d is an Ind-algebraic stack. By definition, we have a 2-Cartesian diagram

$$(5.3.1) \quad \begin{array}{ccc} \mathcal{W}_{d,h} & \longrightarrow & \mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h} \\ \downarrow & & \downarrow \\ \mathcal{X}_d & \longrightarrow & \mathcal{R}_d^{\Gamma_{\mathrm{disc}}} \end{array}$$

If $h' \geq h$ then the closed immersion $\mathcal{C}_{d,h} \hookrightarrow \mathcal{C}_{d,h'}$ is compatible with the morphisms from each of its source and target to \mathcal{R}_d , and so we obtain a closed immersion

$$(5.3.2) \quad \mathcal{W}_{d,h} \hookrightarrow \mathcal{W}_{d,h'}.$$

By construction, the morphisms $\mathcal{W}_{d,h} \rightarrow \mathcal{X}_d$ are compatible, as h varies, with the closed immersions (5.3.2). Thus we also obtain a morphism

$$(5.3.3) \quad \varinjlim_h \mathcal{W}_{d,h} \rightarrow \mathcal{X}_d.$$

Roughly speaking, we will prove that \mathcal{X}_d is an Ind-algebraic stack by showing that it is the ‘‘scheme-theoretic image’’ of the morphism $\varinjlim_h \mathcal{W}_{d,h} \rightarrow \mathcal{R}_d^{\Gamma_{\mathrm{disc}}}$ induced by (5.3.3). More precisely, choose $s \geq 0$, and consider the composite

$$(5.3.4) \quad \mathcal{W}_{d,h,s} \rightarrow \mathcal{W}_{d,h} \rightarrow \mathcal{X}_d \rightarrow \mathcal{R}_d^{\Gamma_{\mathrm{disc}}}.$$

This admits the alternative factorization

$$\mathcal{W}_{d,h,s} \rightarrow \mathcal{W}_{d,h} \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}} \times_{\mathcal{R}_d} \mathcal{C}_{d,h} \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}}.$$

Proposition 5.2.3 shows that the composite of the first two arrows is a closed embedding of finite presentation, while Corollary 4.2.3 shows that the third arrow is representable by algebraic spaces, proper, and of finite presentation. Thus (5.3.4) is representable by algebraic spaces, proper, and of finite presentation.

Fix an integer $a \geq 1$, and write $\mathcal{W}_{d,h,s}^a := \mathcal{W}_{d,h,s} \times_{\text{Spf } \mathbf{Z}_p} \text{Spec } \mathbf{Z}/p^a$. Proposition 5.2.3 shows that $\mathcal{W}_{d,h,s}^a$ is a p -adic formal algebraic stack of finite presentation over $\text{Spf } \mathbf{Z}_p$, and so $\mathcal{W}_{d,h,s}^a$ is an algebraic stack, and a closed substack of $\mathcal{W}_{d,h,s}$.

Note that since at this point we don't know that \mathcal{X}_d is Ind-algebraic, we can't directly define a scheme-theoretic image of $\mathcal{W}_{d,h,s}^a$ in \mathcal{X}_d . It might be possible to do this using the formalism of [EG21]; we take a slightly different approach.

5.3.5. Definition. We let $\mathcal{X}_{d,h,s}^a$ denote the scheme-theoretic image of the composite

$$(5.3.6) \quad \mathcal{W}_{d,h,s}^a \hookrightarrow \mathcal{W}_{d,h,s} \xrightarrow{(5.3.4)} \mathcal{R}_d^{\Gamma_{\text{disc}}}.$$

This is a morphism of Ind-algebraic stacks, and the scheme-theoretic image has the obvious meaning: since $\mathcal{R}_d^{\Gamma_{\text{disc}}}$ is an Ind-algebraic stack, constructed as the 2-colimit of a directed system of algebraic stacks whose transition morphisms are closed immersions, the morphism (5.3.6), whose domain is a quasi-compact algebraic stack, factors through a closed algebraic substack \mathcal{Z} of $\mathcal{R}_d^{\Gamma_{\text{disc}}}$. We then define $\mathcal{X}_{d,h,s}^a$ to be the scheme-theoretic image of $\mathcal{W}_{d,h,s}^a$ in \mathcal{Z} . Then $\mathcal{X}_{d,h,s}^a$ is a closed algebraic substack of $\mathcal{R}_d^{\Gamma_{\text{disc}}}$, and is independent of the choice of \mathcal{Z} .

Our next goal is to prove that $\mathcal{X}_{d,h,s}^a$ is a (necessarily closed) substack of \mathcal{X}_d . Our argument for this is a little indirect. By definition, it is enough to check that if A is a finite type \mathbf{Z}/p^a -algebra, then for any morphism $\text{Spec } A \rightarrow \mathcal{X}_{d,h,s}^a$, the composite morphism $\text{Spec } A \rightarrow \mathcal{X}_{d,h,s}^a \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}}$ factors through \mathcal{X}_d . More concretely, if M denotes the étale φ -module over A , endowed with a Γ_{disc} -action, associated to the given point $\text{Spec } A \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}}$, then we must show that the Γ_{disc} -action on M is continuous. By Lemma 5.1.3, we need to show that M contains a (not necessarily projective) lattice \mathfrak{M} such that for some $s \geq 0$, we have $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$.

To do this, we note that the natural map $A \rightarrow B := \varprojlim_i A_i$ is injective, where A_i runs over the Artinian quotients of A . It is enough to find such a lattice for M_B , and using that lattices of bounded height are uniformly commensurable, it is in fact enough to find a lattice of some fixed height for each M_{A_i} . We are therefore reduced to the following lemma [EG19, Lem. 3.4.8].

5.3.7. Lemma. *Suppose that M is a projective étale φ -module of rank d over a finite type Artinian \mathbf{Z}/p^a -algebra A , and that M is endowed with an action of Γ_{disc} , such that the corresponding morphism $\text{Spec } A \rightarrow \mathcal{R}_d^{\Gamma_{\text{disc}}}$ factors through $\mathcal{X}_{d,h,s}^a$. Then M contains a φ -invariant lattice \mathfrak{M} of T -height $\leq h$, such that $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$.*

Our proof of this lemma is a little involved. We can immediately reduce to the Artin local case, and then to a problem about the universal framed deformation rings of a fixed φ -module equipped with an action of Γ_{disc} . Writing R for this universal deformation ring, we can consider the subfunctor of deformations which admit a lattice \mathfrak{M} of the required type, and it is straightforward to check that this

is representable by a quotient S of R . The statement of the lemma reduces to showing that $\mathrm{Spf} S$ contains the scheme-theoretic image $\mathrm{Spf} T$ of the morphism

$$X := \mathcal{W}_{d,h,s} \times_{\mathcal{R}_d^{\Gamma_{\mathrm{disc}}}} \mathrm{Spf} R \rightarrow \mathrm{Spf} R.$$

We prove this using the following criterion [EG19, Lem. A.30].

5.3.8. Lemma. *Let $R \rightarrow S$ be a continuous surjection of pro-Artinian local rings, and let $X \rightarrow \mathrm{Spf} R$ be a finite type morphism of formal algebraic spaces.*

Make the following assumption: if A is any finite-type Artinian local R -algebra for which the canonical morphism $R \rightarrow A$ factors through a discrete quotient of R , and for which the canonical morphism $X_A \rightarrow \mathrm{Spec} A$ admits a section, then the canonical morphism $R \rightarrow A$ furthermore factors through S .

Then the scheme-theoretic image of $X \rightarrow \mathrm{Spf} R$ is a closed formal subscheme of $\mathrm{Spf} S$.

The result then follows by unwinding the definitions: the point is that admitting a section to X_A in particular gives a morphism $\mathrm{Spec} A \rightarrow \mathcal{W}_{d,h,s}$, and the corresponding weak Wach module is a lattice of the kind being considered, which gives the required factorisation through S .

Finally, we can prove that \mathcal{X}_d is Ind-algebraic.

5.3.9. Proposition. *The canonical morphism $\varinjlim \mathcal{X}_{d,h,s}^a \rightarrow \mathcal{X}_d$ is an isomorphism. Thus \mathcal{X}_d is an Ind-algebraic stack, and may in fact be written as the inductive limit of algebraic stacks of finite presentation, with the transition maps being closed immersions.*

Proof. We just need to show that if $T = \mathrm{Spec} A$ for a Noetherian \mathbf{Z}/p^a -algebra A , then any morphism $T \rightarrow \mathcal{X}_d$ factors through some $\mathcal{X}_{d,h,s}^a$, or equivalently, that the closed immersion

$$(5.3.10) \quad \mathcal{X}_{d,h,s}^a \times_{\mathcal{X}_d} T \rightarrow T$$

is an isomorphism, for some choice of h and s .

If M denotes the étale (φ, Γ) -module corresponding to the morphism $\mathrm{Spec} A \rightarrow \mathcal{X}_d$, then [EG21, Prop. 5.4.7] shows that we may find a scheme-theoretically dominant morphism $\mathrm{Spec} B \rightarrow \mathrm{Spec} A$ such that M_B is free of rank d . (In outline, the proof of that result is to note that \mathfrak{M} is free if and only if $\mathfrak{M}/T\mathfrak{M}$ is free, so that \mathfrak{M} is automatically free over a dense open subset of $\mathrm{Spec} A$, and use Noetherian induction.) If we show that the composite $\mathrm{Spec} B \rightarrow \mathrm{Spec} A \rightarrow \mathcal{X}_d$ factors through $\mathcal{X}_{d,h,s}^a$ for some h and s , then we see that the morphism $\mathrm{Spec} B \rightarrow \mathrm{Spec} A$ factors through the closed subscheme $\mathcal{X}_{d,h,s}^a \times_{\mathcal{X}_d} \mathrm{Spec} A$ of $\mathrm{Spec} A$. Since $\mathrm{Spec} B \rightarrow \mathrm{Spec} A$ is scheme-theoretically dominant, this implies that (5.3.10) is indeed an isomorphism, as required.

Since M_B is free, we may choose a φ -invariant free lattice $\mathfrak{M} \subseteq M_B$, of height $\leq h$ for some sufficiently large value of h . Since the Γ_{disc} -action on M , and hence on M_B , is continuous by assumption, Lemma 5.1.3 then shows that $(\gamma^{p^s} - 1)(\mathfrak{M}) \subseteq T\mathfrak{M}$ for some sufficiently large value of s . Then \mathfrak{M} gives rise to a B -valued point of $\mathcal{W}_{d,h,s}^a$, whose image in $\mathcal{R}_d^{\Gamma_{\mathrm{disc}}}$ is equal to the étale φ -module M_B . Thus the morphism $\mathrm{Spec} B \rightarrow \mathcal{X}_d$ corresponding to M_B does indeed factor through $\mathcal{X}_{d,h,s}^a$. \square

LECTURE 6. CRYSTALLINE AND SEMISTABLE MODULI STACKS

From now on we allow K/\mathbf{Q}_p to be arbitrary.

6.1. Breuil–Kisin–Fargues modules admitting all descents. The connection between Breuil–Kisin modules and crystalline representations was first established by Kisin in [Kis06], where he showed that if K_∞ is as in Example 3.1.2, and $\rho : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{Z}}_p)$ is a lattice in a crystalline representation with Hodge–Tate weights in $[0, h]$, then the étale φ -module corresponding to $\rho|_{G_{K_\infty}}$ (via the functors explained in Lecture 3.3) arises from a (unique) Breuil–Kisin module of height at most h . Kisin also showed that on crystalline representations, the restriction from G_K to G_{K_∞} is fully faithful. Unfortunately, when $h > 1$, Kisin’s description of the Breuil–Kisin modules which arise from lattices in crystalline representations is transcendental in nature (it is phrased in terms of a certain meromorphic connection, which only exists after an extension of scalars, admitting at worst log poles), and it is unclear (to us) how to define crystalline moduli stacks purely in terms of Breuil–Kisin modules for K_∞ .

Our approach to constructing crystalline and semistable moduli stacks relies on the observation that while most Breuil–Kisin modules do not give crystalline or semistable representations of G_K , this is “mostly” for the simple reason that the corresponding G_{K_∞} -representations do not even extend to G_K . However, if ρ is a G_K -representation such that $\rho|_{G_{K_\infty}}$ admits a Breuil–Kisin module, then ρ is “almost” semistable: in particular, ρ is potentially semistable, and becomes semistable over $K(\pi^{1/p^m})$ for some m depending only on K . Using this result it is straightforward to deduce that ρ is semistable if and only if $\rho|_{G_{K_\infty}}$ admits a Breuil–Kisin module for all possible choices of $(\pi \text{ and } (\pi^{1/p^n})_{n \geq 0})$. This is the main idea behind our construction of the stacks.

In fact, this statement (that $\rho|_{G_{K_\infty}}$ admits a Breuil–Kisin module only if ρ is potentially semistable) was only proved after [EG19] was completed, by Hui Gao [Gao19]. Instead, we use a slightly weaker statement, which is proved by TG and Tong Liu in [EG19, App. D] (see Theorem 6.1.5 below).

For each choice of uniformiser π of K , and each choice $\pi^b \in \mathcal{O}_{\mathcal{C}}^b$ of p -power roots of π , we write \mathfrak{S}_{π^b} for $\mathfrak{S} = W(k)[[u]]$, regarded as a subring of $\mathbf{A}_{\mathrm{inf}}$ via $u \mapsto [\pi^b]$. For each choice of π^b and each $s \geq 0$ we write $K_{\pi^b, s}$ for $K(\pi^{1/p^s})$, and $K_{\pi^b, \infty}$ for $\cup_s K_{\pi^b, s}$. Write $E_\pi(u)$ for the Eisenstein polynomial for π , and E_{π^b} for its image in $\mathbf{A}_{\mathrm{inf}}$. There is a natural ring homomorphism $\theta : \mathbf{A}_{\mathrm{inf}} \rightarrow \mathcal{O}_{\mathcal{C}}$, which satisfies $\theta([x]) = x^\sharp$ for any $x \in \mathcal{O}_{\mathcal{C}}^b$; and E_{π^b} is a generator of $\ker \theta$.

6.1.1. Definition. A Breuil–Kisin–Fargues G_K -module of height at most h is a Breuil–Kisin–Fargues module of height at most h which is equipped with a semilinear G_K -action which commutes with φ .

6.1.2. Remark. Note that if $\mathfrak{M}^{\mathrm{inf}}$ is a Breuil–Kisin–Fargues G_K -module, then $W(\mathbf{C}^b) \otimes_{\mathbf{A}_{\mathrm{inf}}} \mathfrak{M}^{\mathrm{inf}}$ is naturally a (G_K, φ) -module in the sense of Definition 3.4.2.

6.1.3. Definition. Let $\mathfrak{M}^{\mathrm{inf}}$ be a Breuil–Kisin–Fargues G_K -module of height at most h with a semilinear G_K -action. Then we say that $\mathfrak{M}^{\mathrm{inf}}$ admits all descents if the following conditions hold.

- (1) For every choice of π and π^b , there is a Breuil–Kisin module \mathfrak{M}_{π^b} of height at most h with $\mathfrak{M}_{\pi^b} \subset (\mathfrak{M}^{\mathrm{inf}})^{G_K}_{\pi^b, \infty}$ for which the induced morphism $\mathbf{A}_{\mathrm{inf}} \otimes_{\mathfrak{S}_{\pi^b}} \mathfrak{M}_{\pi^b} \rightarrow \mathfrak{M}^{\mathrm{inf}}$ is an isomorphism.
- (2) The $W(k)$ -submodule $\mathfrak{M}_{\pi^b}/[\pi^b]\mathfrak{M}_{\pi^b}$ of $W(\overline{k}) \otimes_{\mathbf{A}_{\mathrm{inf}}} \mathfrak{M}^{\mathrm{inf}}$ is independent of the choice of π and π^b .

- (3) The \mathcal{O}_K -submodule $\varphi^*\mathfrak{M}_{\pi^b}/E_{\pi^b}\varphi^*\mathfrak{M}_{\pi^b}$ of $\mathcal{O}_{\mathbf{C} \otimes_{\theta, \mathbf{A}_{\text{inf}}}}\varphi^*\mathfrak{M}^{\text{inf}}$ is independent of the choice of π and π^b .

6.1.4. Definition. Let $\mathfrak{M}^{\text{inf}}$ be a Breuil–Kisin–Fargues G_K -module which admits all descents. We say that $\mathfrak{M}^{\text{inf}}$ is furthermore *crystalline* if for each choice of π and π^b , and each $g \in G_K$, we have

$$(g-1)(\mathfrak{M}_{\pi^b}) \subset \varphi^{-1}([\varepsilon]-1)[\pi^b]\mathfrak{M}^{\text{inf}}.$$

There is an equivalence of categories between the category of (G_K, φ) -modules M of rank d and the category of free \mathbf{Z}_p -modules T of rank d which are equipped with a continuous action of G_K , with the Galois representation corresponding to M being given by $T(M) = M^{\varphi=1}$. Write $V(M) := T(M) \otimes_{\mathbf{Z}_p} \mathbf{Q}_p$.

We deduce the following theorem [EG19, Thm. F.11] from the results of Tong Liu’s paper [Liu18] and Fargues’ correspondence between Breuil–Kisin–Fargues modules and \mathbf{B}_{dR}^+ -lattices [BMS18, Thm. 4.28].

6.1.5. Theorem (T.G. and Tong Liu). *Let M be a (G_K, φ) -module. Then $V(M)$ is semistable with Hodge–Tate weights in $[0, h]$ if and only if there is a (necessarily unique) Breuil–Kisin–Fargues G_K -module $\mathfrak{M}^{\text{inf}}$ which is of height at most h , which admits all descents, and which satisfies $M = W(\mathbf{C}^b) \otimes_{\mathbf{A}_{\text{inf}}} \mathfrak{M}^{\text{inf}}$.*

Furthermore, $V(M)$ is crystalline if and only if $\mathfrak{M}^{\text{inf}}$ is crystalline.

Very recently, Bhargav Bhatt and Peter Scholze have announced a classification of lattices in crystalline G_K -representations in terms of prismatic F -crystals; this presumably implies Theorem 6.1.5 (in the crystalline case), and puts it in its natural context.

6.2. Definition of the crystalline stacks. For simplicity of exposition, we will only address the crystalline case from now on; the proofs in the semistable case are identical (in fact slightly simpler, because the crystalline case has the extra condition on the Breuil–Kisin–Fargues modules). The actual definition of these stacks, and the proof of their basic properties, is quite involved, but the basic idea is quite simple: we will take the scheme-theoretic image in \mathcal{X}_d of the moduli stack of Breuil–Kisin–Fargues G_K -modules which admit all descents.

Firstly, note that Definition 6.1.4 carries over in an obvious fashion to the case with coefficients in a topologically of finite type p -adically complete \mathbf{Z}_p -algebra A (see [EG19, Defn. 4.2.4]). Using the faithful flatness results of Proposition 3.2.3, one can check that the existence of a descent to a Breuil–Kisin module depends only on π , and not on π^b , and that any such descent is unique ([EG19, Lem. 4.2.7, 4.2.8]).

6.2.1. Definition. For any $h \geq 0$ we let $\mathcal{C}_{d, \text{crys}, h}^a$ denote the limit preserving category of groupoids over $\text{Spec } \mathbf{Z}/p^a$ determined by decreeing, for any finite type \mathbf{Z}/p^a -algebra A , that $\mathcal{C}_{d, \text{crys}, h}^a(A)$ is the groupoid of Breuil–Kisin–Fargues G_K -modules with A -coefficients, which are of height at most h , which admit all descents, and which are crystalline. We let $\mathcal{C}_{d, \text{crys}, h} := \varinjlim_a \mathcal{C}_{d, \text{crys}, h}^a$.

It is easy to check that if A is a p -adically complete \mathbf{Z}_p -algebra which is topologically of finite type, then $\mathcal{C}_{d, \text{crys}, h}(A)$ is the groupoid of Breuil–Kisin–Fargues G_K -modules with A -coefficients, which are of height at most h , which admit all descents, and which are crystalline. There is a natural morphism

$$\mathcal{C}_{d, \text{crys}, h} \rightarrow \mathcal{X}_d,$$

which is defined, for finite type \mathbf{Z}/p^a -algebras, via $\mathfrak{M}^{\text{inf}} \mapsto W(\mathbf{C}^\flat)_A \otimes_{\mathbf{A}_{\text{inf},A}} \mathfrak{M}^{\text{inf}}$, with the target object being regarded as an A -valued point of \mathcal{X}_d via the equivalence of Proposition 3.4.3.

For now we will admit the following theorem [EG19, Thm. 4.5.18], which we prove in the following section.

6.2.2. Theorem. *$\mathcal{C}_{d,\text{crys},h}$ is a p -adic formal algebraic stack of finite presentation and affine diagonal. The morphism $\mathcal{C}_{d,\text{crys},h} \rightarrow \mathcal{X}_d$ is representable by algebraic spaces, proper, and of finite presentation.*

Let $\mathcal{C}_{d,\text{crys},h}^{\text{fl}}$ denote the flat part of $\mathcal{C}_{d,\text{crys},h}$ (i.e. the maximal substack which is flat over $\text{Spf } \mathbf{Z}_p$; see [Eme, Ex. 9.11]). Then we define $\mathcal{X}_d^{\text{crys},h}$ to be the scheme-theoretic image of the morphism $\mathcal{C}_{d,\text{crys},h}^{\text{fl}} \rightarrow \mathcal{X}_d$. The following result is essentially [EG19, Thm. 4.8.12] (for the purposes of exposition, we are ignoring inertial types and Hodge–Tate weights for now).

6.2.3. Theorem. *The closed substack $\mathcal{X}_d^{\text{crys},h}$ of \mathcal{X}_d is a p -adic formal algebraic stack, which is of finite type and flat over $\text{Spf } \mathbf{Z}_p$, and is uniquely determined as a \mathbf{Z}_p -flat closed substack of \mathcal{X}_d by the following property: if A° is a finite flat \mathbf{Z}_p -algebra, then $\mathcal{X}_d^{\text{crys},h}(A^\circ)$ is precisely the subgroupoid of $\mathcal{X}_d(A^\circ)$ consisting of G_K -representations which are crystalline with Hodge–Tate weights contained in $[0, h]$.*

That $\mathcal{X}_d^{\text{crys},h}$ is a p -adic formal algebraic stack and is of finite type and flat over $\text{Spf } \mathbf{Z}_p$ follows from its construction and [EG19, Prop. A.21] (that is, “the scheme-theoretic image of a p -adic formal algebraic stack is a p -adic formal algebraic stack”). The characterisation of its A° -points can be reduced to the local case, and thus to the following important statement about versal rings: fix a point $\text{Spec } \mathbf{F} \rightarrow \mathcal{X}_d(\mathbf{F})$ for some finite field \mathbf{F} , giving rise to a continuous representation $\bar{\rho} : G_K \rightarrow \text{GL}_d(\mathbf{F})$. Let $R_{\bar{\rho}}^{\text{crys},h}$ denote the universal framed deformation $W(\mathbf{F})$ -algebra for lifts of $\bar{\rho}$ which are crystalline with Hodge–Tate weights in $[0, h]$ (which exists by [Kis08]). Then there is an induced morphism $\text{Spf } R_{\bar{\rho}}^{\text{crys},h} \rightarrow \mathcal{X}_d$, and we have [EG19, Prop. 4.8.10]:

6.2.4. Proposition. *The morphism $\text{Spf } R_{\bar{\rho}}^{\text{crys},h} \rightarrow \mathcal{X}_d$ factors through a versal morphism $\text{Spf } R_{\bar{\rho}}^{\text{crys},h} \rightarrow \mathcal{X}_d^{\text{crys},h}$.*

Proof. This is proved using Lemma 5.3.8, in a similar way to the way we used it in the sketch proof of Lemma 5.3.7; the actual argument is slightly more complicated as we prove an algebraization statement and then work with p inverted. \square

6.3. Canonical actions. We now prove Theorem 6.2.2. We will deduce it from the corresponding statements for the stacks $\mathcal{C}_{d,h}$ of Breuil–Kisin modules, which we discussed in Lecture 4; the key point is (roughly) to show that the additional structure of a G_K -action admitting all descents is only a finite amount of data.

The key finiteness property comes from the following reinterpretation of some of the arguments of [CL11, §2] (see [EG19, §4.3, 4.4]).

6.3.1. Proposition. *For any fixed a, h , and any sufficiently large N , there is a positive integer $s(a, h, N)$ with the property that for any finite type \mathbf{Z}/p^a -algebra A , any projective Breuil–Kisin module \mathfrak{M} of height at most h , and any $s \geq s(a, h, N)$, there is a unique continuous action of $G_{K_s} = G_{K(\pi^{1/p^s})}$ on $\mathfrak{M}^{\text{inf}} := \mathbf{A}_{\text{inf},A} \otimes_{\mathfrak{S}_A} \mathfrak{M}$*

which commutes with φ and is semi-linear with respect to the natural action of G_{K_s} on $\mathbf{A}_{\text{inf}, A}$, with the additional property that for all $g \in G_{K_s}$ we have $(g-1)(\mathfrak{M}) \subset u^N \mathfrak{M}^{\text{inf}}$.

Conversely, if $\mathfrak{M}^{\text{inf}}$ be a Breuil–Kisin–Fargues G_K -module of height at most h with A -coefficients, which admits all descents, then for each choice of π^b , and each $s \geq s(a, h, N)$, the action of $G_{K_{\pi^b, s}}$ agrees with the canonical action obtained from \mathfrak{M}_{π^b} as above.

The proof of the first part is quite straightforward; the point is that by a standard argument with the φ -structure, one can uniquely upgrade approximate homomorphisms (i.e. maps which are homomorphisms modulo u^N) between Breuil–Kisin(–Fargues) modules to actual homomorphisms (see [EG19, Lem. 4.3.2]). We can write down an “approximate action” of G_{K_s} on $\mathfrak{M}^{\text{inf}}$, by just letting it act trivially on \mathfrak{M} (after fixing some basis - since the action is semilinear, it doesn’t literally make sense to ask that it be trivial).

For the second part, it’s easy to check that this holds for $s \gg 0$ just by continuity of the action of G_K on $\mathfrak{M}^{\text{inf}}$, and to deduce it for the optimal value of s , we consider the subgroup of G_K generated by the $G_{K_{\pi^b, s}}$ as we run over all choices of π^b .

Then for any $h \geq 0$ and any choice of π^b , we write $\mathcal{C}_{\pi^b, d, h}$ for the moduli stack of rank d Breuil–Kisin modules for $\mathfrak{S}_{\pi^b, A}$ of height at most h , and $\mathcal{R}_{\pi^b, d}$ for the corresponding stack of étale φ -modules. We have a natural morphism $\mathcal{X}_{K, d} \rightarrow \mathcal{R}_{\pi^b, d}$ given by Proposition 3.4.3 and restriction from G_K to G_{K_∞} ; for each $s \geq 1$, we can factor this as

$$\mathcal{X}_{K, d} \rightarrow \mathcal{X}_{K_{\pi^b, s}, d} \rightarrow \mathcal{R}_{\pi^b, d}.$$

Then for any fixed a, h , and any N and $s \geq s(a, h, N)$ as in Proposition 6.3.1, there is a canonical morphism $\mathcal{C}_{\pi^b, d, h}^a \rightarrow \mathcal{X}_{K_{\pi^b, s}, d}^a$ obtained from the canonical action of Proposition 6.3.1, which fits into a commutative triangle

$$\begin{array}{ccc} & \mathcal{C}_{\pi^b, d, h}^a & \\ & \swarrow & \downarrow \\ \mathcal{X}_{K_{\pi^b, s}, d}^a & \longrightarrow & \mathcal{R}_{\pi^b, d}^a \end{array}$$

We then define a stack $\mathcal{C}_{\pi^b, s, d, h}^a$ by the requirement that it fits into a 2-Cartesian diagram

$$(6.3.2) \quad \begin{array}{ccc} \mathcal{C}_{\pi^b, s, d, h}^a & \longrightarrow & \mathcal{C}_{\pi^b, d, h}^a \\ \downarrow & & \downarrow \\ \mathcal{X}_{K, d}^a & \longrightarrow & \mathcal{X}_{K_{\pi^b, s}, d}^a \end{array}$$

The lower horizontal arrow in this diagram is again defined via Proposition 3.4.3, and it is not so hard to show (using the properties that we have proved about $\mathcal{X}_{K, d}$ and its diagonal) that it is representable by algebraic spaces and of finite presentation. Since $\mathcal{C}_{\pi^b, d, h}^a$ is an algebraic stack of finite presentation over \mathbf{Z}/p^a , it follows that $\mathcal{C}_{\pi^b, s, d, h}^a$ is also an algebraic stack of finite presentation over \mathbf{Z}/p^a . It is also

straightforward to check that the right hand vertical arrow in this diagram is representable by algebraic spaces, proper, and of finite presentation, so that the same is true of the left hand vertical arrow.

Given all this, the key statement remaining to be proved is the following [EG19, Prop. 4.5.15].

6.3.3. Proposition. *For each π^b and each s as above, there is a natural closed immersion $\mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{C}_{\pi^b,s,d,h}^a$. In particular, $\mathcal{C}_{d,\text{crys},h}^a$ is an algebraic stack of finite presentation over \mathbf{Z}/p^a , and its diagonal is affine.*

Proof. We have a morphism $\mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{X}_{K,d}^a$ (given by extending scalars to $W(\mathbf{C}^b)$), and it follows that there is a natural morphism $\mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{C}_{\pi^b,d,h}^a$, defined via $\mathfrak{M}^{\text{inf}} \mapsto \mathfrak{M}_{\pi^b}$. The composite morphisms $\mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{C}_{\pi^b,d,h}^a \rightarrow \mathcal{X}_{K_{\pi^b,s},d}^a$ and $\mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{X}_{K,d}^a \rightarrow \mathcal{X}_{K_{\pi^b,s},d}^a$ coincide by construction, so we have an induced morphism

$$(6.3.4) \quad \mathcal{C}_{d,\text{crys},h}^a \rightarrow \mathcal{C}_{\pi^b,s,d,h}^a$$

which we need to show is a closed immersion.

To see that (6.3.4) is at least a monomorphism, it is enough to note that if A is a finite type \mathbf{Z}/p^a -algebra, and $\mathfrak{M}^{\text{inf}}$ is a Breuil–Kisin–Fargues module over A which admits all descents, then $\mathfrak{M}^{\text{inf}} = \mathbf{A}_{\text{inf},A} \otimes_{\mathfrak{S}_{\pi^b,A}} \mathfrak{M}_{\pi^b}$ is determined by \mathfrak{M}_{π^b} , and the G_K -action on $\mathfrak{M}^{\text{inf}}$ is determined by the G_K -action on $W(\mathbf{C}^b)_A \otimes_{\mathbf{A}_{\text{inf},A}} \mathfrak{M}^{\text{inf}}$.

Finally, the proof (6.3.4) is a closed immersion is a bit more involved: we need to show that the conditions that $\mathfrak{M}^{\text{inf}} = \mathbf{A}_{\text{inf},A} \otimes_{\mathfrak{S}_{\pi^b,A}} \mathfrak{M}_{\pi^b}$ is G_K -stable and admits all descents are closed conditions. We make repeated use of the results of [EG19, App. B], particularly [EG19, Lem. B.28] to show that the vanishing loci of various morphisms of Breuil–Kisin–Fargues modules are closed. \square

6.4. Potentially crystalline moduli stacks with fixed Hodge–Tate weights.

For applications, it is important to be able to fix the Hodge–Tate weights in our crystalline moduli stacks, and also to consider representations which are only potentially crystalline. To this end, let L/K be a fixed finite Galois extension; then there is an obvious notion of a Breuil–Kisin–Fargues G_K -module which admits all descents over L , and an obvious extension of the above results to give stacks of G_K -representations which become crystalline over L with Hodge–Tate weights contained in $[0, h]$. We write $\mathcal{C}_{d,\text{crys},h}^{L/K,\square}$ for the corresponding stack of Breuil–Kisin–Fargues modules.

Since the inertia types and Hodge–Tate weights are discrete invariants, we at least morally expect these stacks to decompose as a disjoint union of stacks of potentially crystalline representations with fixed inertial and Hodge types; so the key point should be to see how to read this data off from the Breuil–Kisin–Fargues modules.

Fix a finite extension E/\mathbf{Q}_p with ring of integers \mathcal{O} , uniformizer ϖ , and residue field \mathbf{F} , and assume that E is large enough to contain the images of all embeddings $K \hookrightarrow \overline{\mathbf{Q}_p}$. We will abusively write \mathcal{X}_d for the base change of \mathcal{X}_d to $\text{Spf } \mathcal{O}$, without further comment. From now on A° will denote a p -adically complete flat \mathcal{O} -algebra which is topologically of finite type over \mathcal{O} , and we write $A = A^\circ[1/p]$.

Let \mathfrak{M}_{A° be a Breuil–Kisin–Fargues G_K -module with A° -coefficients which admits all descents over L , and write $\overline{\mathfrak{M}}_{A^\circ} := \mathfrak{M}_{A^\circ, \pi^b} / [\pi^b] \mathfrak{M}_{A^\circ, \pi^b}$ (for some choice

of π^b , with π now denoting a uniformiser of L). Then $\overline{\mathfrak{M}}_{A^\circ}$ has a natural $W(l) \otimes_{\mathbf{Z}_p} A$ -semilinear action of $\text{Gal}(L/K)$, which is defined as follows: if $g \in \text{Gal}(L/K)$, then $g(\mathfrak{M}_{A^\circ, \pi^b}) = \mathfrak{M}_{A^\circ, g(\pi^b)}$, so the morphism $g : \mathfrak{M}_{A^\circ, \pi^b} \rightarrow g(\mathfrak{M}_{A^\circ, \pi^b}) = \mathfrak{M}_{A^\circ, g(\pi^b)}$ induces a morphism

$$g : \mathfrak{M}_{A^\circ, \pi^b} / [\pi^b] \mathfrak{M}_{A^\circ, \pi^b} \rightarrow \mathfrak{M}_{A^\circ, g(\pi^b)} / [g(\pi^b)] \mathfrak{M}_{A^\circ, g(\pi^b)},$$

and the source and target are both canonically identified with $\overline{\mathfrak{M}}_{A^\circ}$.

This action of $\text{Gal}(L/K)$ on $\overline{\mathfrak{M}}_{A^\circ}$ induces an $L_0 \otimes_{\mathbf{Q}_p} A$ -linear action of $I_{L/K}$ on the projective $L_0 \otimes_{\mathbf{Q}_p} A$ -module $\overline{\mathfrak{M}}_{A^\circ} \otimes_{A^\circ} A$. Fix a choice of embedding $\sigma : L_0 \hookrightarrow E$, and let $e_\sigma \in L_0 \otimes_{\mathbf{Q}_p} E$ be the corresponding idempotent. Then $e_\sigma(\overline{\mathfrak{M}}_{A^\circ} \otimes_{A^\circ} A)$ is a projective A -module of rank d , with an A -linear action of $I_{L/K}$. Up to canonical isomorphism, this module does not depend on the choice of σ . We write

$$\text{WD}(\mathfrak{M}_{A^\circ}^{\text{inf}}) := e_\sigma(\overline{\mathfrak{M}}_{A^\circ} \otimes_{A^\circ} A),$$

a projective A -module of rank d with an A -linear action of $I_{L/K}$. It is easy to check that this is compatible with base change (of p -adically complete flat \mathcal{O} -algebras which are topologically of finite type over \mathcal{O}).

We now turn to Hodge types. Fix some choice of π^b a uniformiser of L , write \mathfrak{M}_{A° for $\mathfrak{M}_{\pi^b, A^\circ}$, and u for $[\pi^b]$. For each $0 \leq i \leq h$ we define $\text{Fil}^i \varphi^* \mathfrak{M}_{A^\circ} = \Phi_{\mathfrak{M}_{A^\circ}}^{-1}(E(u)^i \mathfrak{M}_{A^\circ})$, and we set $\text{Fil}^i \varphi^* \mathfrak{M}_{A^\circ} = \varphi^* \mathfrak{M}_{A^\circ}$ for $i < 0$. Then for each $0 \leq i \leq h$,

$$(\text{Fil}^i \varphi^* \mathfrak{M}_{A^\circ} / E(u) \text{Fil}^{i-1} \varphi^* \mathfrak{M}_{A^\circ}) \otimes_{A^\circ} A$$

is a finite projective $L \otimes_{\mathbf{Q}_p} A$ -module, whose formation is compatible with base change (see [EG19, Prop. 4.7.1]). There is a natural action of $\text{Gal}(L/K)$, which is semi-linear with respect to the action of $\text{Gal}(L/K)$ on $L \otimes_{\mathbf{Q}_p} A$ induced by its action on the first factor. Since L/K is a Galois extension, the tensor product $L \otimes_{\mathbf{Q}_p} A$ is an étale $\text{Gal}(L/K)$ -extension of $K \otimes_{\mathbf{Q}_p} A$, and so étale descent allows us to descend $(\varphi^* \mathfrak{M}_{A^\circ} / E(u) \varphi^* \mathfrak{M}_{A^\circ}) \otimes_{A^\circ} A$ to a filtered module over $K \otimes_{\mathbf{Q}_p} A$; concretely, this descent is achieved by taking $\text{Gal}(L/K)$ -invariants. This leads to the following definition.

6.4.1. Definition. In the preceding situation, we write

$$D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}) := ((\varphi^* \mathfrak{M}_{A^\circ} / E(u) \varphi^* \mathfrak{M}_{A^\circ}) \otimes_{A^\circ} A)^{\text{Gal}(L/K)},$$

and more generally, for each $i \geq 0$, we write

$$\text{Fil}^i D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}) := ((\text{Fil}^i \varphi^* \mathfrak{M}_{A^\circ} / E(u) \text{Fil}^{i-1} \varphi^* \mathfrak{M}_{A^\circ}) \otimes_{A^\circ} A)^{\text{Gal}(L/K)}$$

(and for $i < 0$, we write $\text{Fil}^i D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}) := D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}})$). The property of being a finite rank projective module is preserved under étale descent, and so we find that $D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}})$ is a rank d projective $K \otimes_{\mathbf{Q}_p} A$ -module, filtered by projective submodules.

Since A is an E -algebra, we have the product decomposition $K \otimes_{\mathbf{Q}_p} A \xrightarrow{\sim} \prod_{\sigma: K \hookrightarrow E} A$, and so, if we write e_σ for the idempotent corresponding to the factor labeled by σ in this decomposition, we find that

$$D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}) = \prod_{\sigma: K \hookrightarrow E} e_\sigma D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}),$$

where each $e_\sigma D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}})$ is a projective A -module of rank d . For each i , we write

$$\text{Fil}^i e_\sigma D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}) = e_\sigma \text{Fil}^i D_{\text{dR}}(\mathfrak{M}_{A^\circ}^{\text{inf}}).$$

Each $\mathrm{Fil}^i e_\sigma D_{\mathrm{dR}}(\mathfrak{M}_{A^\circ}^{\mathrm{inf}})$ is again a projective A -module.

6.4.2. Definition. A Hodge type $\underline{\lambda}$ of rank d is by definition a set of tuples of integers $\{\lambda_{\sigma,i}\}_{\sigma:K\hookrightarrow\overline{\mathbf{Q}}_p, 1\leq i\leq d}$ with $\lambda_{\sigma,i}\geq\lambda_{\sigma,i+1}$ for all σ and all $1\leq i\leq d-1$. We say that it is *effective* if all $\lambda_{\sigma,i}\geq 0$, and *bounded by h* if all $\lambda_{\sigma,i}\leq h$.

If $\underline{D} := (D_\sigma)_{\sigma:K\hookrightarrow E}$ is a collection of rank d vector bundles over $\mathrm{Spec} A$, labeled (as indicated) by the embeddings $\sigma:K\hookrightarrow E$, then we say that \underline{D} has Hodge type $\underline{\lambda}$ if $\mathrm{Fil}^i D_\sigma$ has constant rank equal to $\#\{j\mid\lambda_{\sigma|K,j}\geq i\}$.

As the notation suggests, the Hodge type of $V(\mathfrak{M}^{\mathrm{inf}})$ agrees with the Hodge type of $D_{\mathrm{dR}}(\mathfrak{M}^{\mathrm{inf}})$. Putting this all together, we have [EG19, Prop. 4.8.2]:

6.4.3. Proposition. *Let L/K be a finite Galois extension. Then the stack $\mathcal{C}_{d,\mathrm{crys},h}^{L/K,\mathrm{fl}}$ is a scheme-theoretic union of closed substacks $\mathcal{C}_{d,\mathrm{crys},h}^{L/K,\mathrm{fl},\underline{\lambda},\tau}$, where $\underline{\lambda}$ runs over all effective Hodge types that are bounded by h , and τ runs over all d -dimensional E -representations of $I_{L/K}$. These latter closed substacks are uniquely characterised by the following property: if A° is a finite flat \mathcal{O} -algebra, then an A° -point of $\mathcal{C}_{d,\mathrm{crys},h}^{L/K,\mathrm{fl}}$ is a point of $\mathcal{C}_{d,\mathrm{crys},h}^{L/K,\mathrm{fl},\underline{\lambda},\tau}$ if and only if the corresponding Breuil–Kisin–Fargues module $\mathfrak{M}_{A^\circ}^{\mathrm{inf}}$ has Hodge type $\underline{\lambda}$ and inertial type τ .*

6.4.4. Remark. It is not obvious (at least to us) from the definition that the Hodge filtration on $D_{\mathrm{dR}}(\mathfrak{M}^{\mathrm{inf}})$ is independent of the choice of π^b ; rather, we deduce this independence from its compatibility with the Hodge filtration on $D_{\mathrm{dR}}(V(\mathfrak{M}^{\mathrm{inf}}))$.

We then define the corresponding stacks $\mathcal{X}_d^{\mathrm{crys},\underline{\lambda},\tau}$; we can extend the definition to possibly negative Hodge–Tate weights by twisting by an appropriate power of the cyclotomic character. Finally we can deduce [EG19, Thm. 4.8.14]:

6.4.5. Theorem. *The algebraic stack $\mathcal{X}_d^{\mathrm{crys},\underline{\lambda},\tau}\times_{\mathrm{Spf}\mathcal{O}}\mathrm{Spec}\mathbf{F}$ is equidimensional of dimension*

$$\sum_{\sigma}\#\{1\leq i<j\leq d\mid\lambda_{\sigma,i}>\lambda_{\sigma,j}\}.$$

In particular, if $\underline{\lambda}$ is regular, then the algebraic stack $\mathcal{X}_d^{\mathrm{crys},\underline{\lambda},\tau}\times_{\mathrm{Spf}\mathcal{O}}\mathrm{Spec}\mathbf{F}$ is equidimensional of dimension $[K:\mathbf{Q}_p]d(d-1)/2$.

This follows from Proposition 6.2.4, from which it follows that the versal rings to $\mathcal{X}_d^{\mathrm{crys},\underline{\lambda},\tau}$ are given by the corresponding deformation rings $R_{\overline{\rho}}^{\mathrm{crys},\underline{\lambda},\tau}$, and the computation of the dimensions of these deformation rings in [Kis08] (which comes down to a computation on the generic fibre, using weakly admissible modules).

LECTURE 7. THE HERR COMPLEX, AND FAMILIES OF EXTENSIONS

We continue to allow K/\mathbf{Q}_p to be an arbitrary finite extension.

7.1. Herr Complex. Let M be a projective étale (φ, Γ) -module over $\mathbf{A}_{K,A}$, where A can be any \mathbf{Z}_p -algebra, but to make nontrivial statements we will often think about the case where A is of finite type over \mathbf{Z}/p^a for some $a\geq 1$. Then we define the *Herr complex*

$$\mathcal{C}^\bullet(M) = [M \xrightarrow{(\varphi-1, \gamma-1)} M \oplus M \xrightarrow{(\gamma-1)\oplus(1-\varphi)} M]$$

concentrated in degrees 0, 1, 2. Note that this is a complex of A -modules, rather than $\mathbf{A}_{K,A}$ -modules (the terms are $\mathbf{A}_{K,A}$ -modules, but the maps are only $\mathbf{A}_{K,A}$ -semilinear with respect to the φ - or γ -actions). The following is [EG19, Thm. 5.1.21].

7.1.1. Theorem. *Suppose A is a Noetherian \mathbf{Z}/p^a -algebra with A/p countable. Then the Herr complex $\mathcal{C}^\bullet(M)$ is a perfect complex of A -modules. If B is a finite type A -algebra, then there is a natural quasi-isomorphism $\mathcal{C}^\bullet(M) \otimes_A^{\mathbf{L}} B \xrightarrow{\sim} \mathcal{C}^\bullet(M \otimes_{\mathbf{A}_{K,A}} \mathbf{A}_{K,B})$.*

Proof. Because A is Noetherian, $\mathbf{A}_{K,A}$ is flat over A . Since M is a projective $\mathbf{A}_{K,A}$ -module, it is therefore also a flat A -module, so the Herr complex is a complex of flat A -modules (which are however not finite over A , but only over $\mathbf{A}_{K,A}$). To check that $\mathcal{C}^\bullet(M)$ is perfect, it therefore suffices to show that the cohomology A -modules are of finite type over A . By Nakayama's lemma for complexes, it in fact suffices to prove the theorem in the case that A is an \mathbf{F}_p -algebra, so we assume this from now on.

If A is in fact a finite-dimensional \mathbf{F}_p -vector space, Herr [Her98] showed that $\mathcal{C}^\bullet(M)$ computes $H^\bullet(G_K, \rho)$, where ρ is the associated G_K -representation corresponding to M ; so the statement we want to prove is analogous to proving a version of the finiteness of Galois cohomology in the presence of coefficients. Our proof largely follows that of Herr.

We don't have an equivalence between (φ, Γ) -modules and Galois representations for general coefficient rings, but we have similar behavior, as in the following lemma.

7.1.2. Lemma. *$H^i(\mathcal{C}^\bullet(M_1^\vee \otimes M_2)) = \text{Ext}_{(\varphi, \Gamma)/\mathbf{A}_{K,A}}^i(M_1, M_2)$ for $i = 0, 1$. In addition, $H^2(\mathcal{C}^\bullet(\text{ad } M))$ controls the obstructions to infinitesimal deformations of M .*

It may well be the case that $H^2(\mathcal{C}^\bullet(M_1^\vee \otimes M_2)) = \text{Ext}_{(\varphi, \Gamma)/\mathbf{A}_{K,A}}^2(M_1, M_2)$, but we have not attempted to verify this.

There is an A -linear map $\psi : \mathbf{A}_{K,A} \rightarrow \mathbf{A}_{K,A}$, which is given by $\psi = \frac{1}{p} \text{tr } \varphi$, and an induced map $\psi : M \rightarrow M$, which is defined so that for $a \in A$ and $m \in M$ we have $\psi(\varphi(am)) = a\psi(m)$ and $\psi(a\varphi(m)) = \psi(a)m$.

Since φ commutes with Γ , so does ψ , and we have an induced morphism $(1 - \gamma) : M^{\psi=0} \rightarrow M^{\psi=0}$. In fact, this morphism is an isomorphism. The proof of this takes some work, but at least morally it has the following explanation in the simplest case, when $A = \mathbf{F}_p$, $K = \mathbf{Q}_p$ and $M = \mathbf{A}_{\mathbf{Q}_p}$. Working instead with $\tilde{\Gamma}$ and $(\mathbf{A}'_{\mathbf{Q}_p})^+ = \mathbf{Z}_p[[\mathbf{Z}_p]]$, we have $\mathbf{F}_p[[\mathbf{Z}_p]]^{\psi=0} = \mathbf{F}_p[[\mathbf{Z}_p^\times]]$ (as an $\mathbf{F}_p[\tilde{\Gamma}]$ -module) $= \mathbf{F}_p[\Delta] \otimes_{\mathbf{F}_p} \mathbf{F}_p[[x]]$, where $x = \gamma - 1$. Thus we see that $(\gamma - 1)$ acts on $\mathbf{F}_p[[x]]$ by multiplication by x , which is injective. After passing to the fraction field of $\mathbf{F}_p[[\mathbf{Z}_p]]$, i.e. after passing to $\mathbf{A}'_{\mathbf{Q}_p}$, i.e. after inverting $T'_{\mathbf{Q}_p}$, one finds, upon passing to ψ -invariants, that this corresponds to inverting x ; i.e. that $(\mathbf{A}'_{\mathbf{Q}_p})^{\psi=0} = \mathbf{F}_p[[\Delta] \times \mathbf{F}_p((x))$. Thus multiplication by $\gamma - 1 = x$ now acts invertibly.

Using the ψ operator, we can give an alternative description of the Herr complex as follows. Let $\mathcal{C}_\psi^\bullet(M)$ be the complex in degrees 0, 1, 2 given by

$$0 \longrightarrow M \xrightarrow{(\psi-1, \gamma-1)} M \oplus M \xrightarrow{(\gamma-1) \oplus (1-\psi)} M \longrightarrow 0.$$

We have a morphism of complexes $\mathcal{C}^\bullet(M) \rightarrow \mathcal{C}_\psi^\bullet(M)$ given by

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \xrightarrow{(\varphi-1, \gamma-1)} & M \oplus M & \xrightarrow{(\gamma-1) \oplus (1-\varphi)} & M & \longrightarrow & 0 \\ & & \downarrow 1 & & \downarrow (-\psi, 1) & & \downarrow -\psi & & \\ 0 & \longrightarrow & M & \xrightarrow{(\psi-1, \gamma-1)} & M \oplus M & \xrightarrow{(\gamma-1) \oplus (1-\psi)} & M & \longrightarrow & 0 \end{array}$$

That this is a morphism of complexes follows from the facts that $\psi \circ \varphi = \text{id}$, and that ψ commutes with γ ; and since $(1 - \gamma) : M^{\psi=0} \rightarrow M^{\psi=0}$ is an isomorphism, we see that is in fact a quasi-isomorphism.

With some more work, we can show that this morphism induces topological isomorphisms of the cohomology groups. We deduce the required finiteness from this: the basic idea is that a Frobenius truncation argument shows that the cohomology groups of $\mathcal{C}^\bullet(M)$ are subquotients of finitely generated $A((T))/A[[T]]$ -modules, and thus have the discrete topology, while the cohomology groups of $\mathcal{C}_\psi^\bullet(M)$ are subquotients of finitely generated $A[[T]]$ -modules (so the slogan is: “discrete and compact implies finite”). It is in these arguments that we use the countability assumption on A , in order to use the theory of Polish groups.

Finally, to see that the obvious map of complexes $\mathcal{C}^\bullet(M) \otimes_A B \xrightarrow{\sim} \mathcal{C}^\bullet(M \otimes_{\mathbf{A}_{K,A}} \mathbf{A}_{K,B})$ induces a quasi-isomorphism, we use an argument of [KPX14] to formally reduce to the case that B is a finite A -module, in which case the claim follows easily as $\mathbf{A}_{K,B} = \mathbf{A}_{K,A} \otimes_A B$. \square

7.2. Constructing families of extensions. Our next goal is to cover the underlying reduced substack $\mathcal{X}_{d,\text{red}}$ by certain families of étale (φ, Γ) -modules, having rather explicit parameterizations, with the goal of showing that $\mathcal{X}_{d,\text{red}}$ is of finite presentation over \mathbf{F}_p , and of bounding its dimension from above,

We begin with the case of $K = \mathbf{Q}_p$ and $d = 1$, and to this end return briefly to the discussions of Lecture 1.4. By local class field theory (see also Lecture 1.3), the $\overline{\mathbf{F}}_p^\times$ -valued characters of G_K are all of the form $\omega^i \text{ur}_\alpha$, for $i = 0, \dots, p-2$ and $\alpha \in \overline{\mathbf{F}}_p^\times$. Going over to the setting of (φ, Γ) -modules, we can take the trivial (φ, Γ) -module over $\mathbf{A}_{\mathbf{Q}_p, \mathbf{F}_p[\alpha, \alpha^{-1}]}$ and twist φ by α and Γ by ω^i , to get a \mathbf{G}_m worth of representations (one for each $\alpha \in \overline{\mathbf{F}}_p^\times$; here i is being kept fixed), but also a \mathbf{G}_m worth of scalar automorphisms. As we run over i we obtain a map

$$\bigsqcup_{i=0}^{p-2} [\mathbf{G}_m / \mathbf{G}_m] \rightarrow \mathcal{X}_{1,\text{red}}$$

which with a little work can be shown to be an isomorphism.

Now let $d = 2$. Any irreducible representation has the form

$$\text{ur}_\alpha \text{Ind}_{G_{\mathbf{Q}_p^2}}^{G_{\mathbf{Q}_p}} \omega_2^i$$

(for some $0 < i < p^2 - 1$ with $(p+1) \nmid i$). This gives us a point $[\mathbf{G}_m / \mathbf{G}_m] \hookrightarrow \mathcal{X}_{2,\text{red}}$ (again the α gives a \mathbf{G}_m and then we quotient by scalars). But we know (looking ahead to Theorem 9.1.1) that $\dim \mathcal{X}_{2,\text{red}} = 1$ and that it’s equidimensional, so the irreducible representations lie in a substack of positive codimension. We can now consider families of reducible representations, so we want to look at representations coming from $\text{Ext}^1(\text{ur}_\alpha \omega^i, \text{ur}_\beta \omega^j)$. For generic choices of α, β there’s a unique non-trivial extension, and the intuition is that as α, β vary, the extension class varies

with them, and then you mod out by scalar endomorphisms to get the expected one-dimensional stack (two dimensions coming from the choice of α, β).

In fact, we want to apply this same technique in more generality. The idea is to begin with some representation, and then to iteratively build its spaces of extensions by some given irreducible representations. More generally, assume that we're given a family $\bar{\rho}_T \rightarrow T$ of rank d étale (φ, Γ) -modules, where T is a reduced finite type variety over $\bar{\mathbf{F}}_p$. (This is our starting representation — of course in practice we work with a (φ, Γ) -module, but we think of it as being a family of Galois representations.) Then take $\bar{\sigma}$ an irreducible Galois representation (or, more properly, the associated étale (φ, Γ) -module). As a technical hypothesis, assume furthermore that the spaces $\text{Ext}_{G_K}^2(\bar{\sigma}, \bar{\rho}_t)$ are of constant dimension when we vary over t over the closed points of T .

The base change property for $\mathcal{C}^\bullet(M)$ in Theorem 7.1.1 involves a derived tensor product, so there's a spectral sequence relating $\mathcal{C}^\bullet(M)$ and $\mathcal{C}^\bullet(M \otimes_A B)$, but since H^2 is the highest degree, it satisfies naive base change. Therefore, our assumption on the Ext^2 having constant degree implies that $H^2(\mathcal{C}^\bullet(\bar{\sigma}^\vee \otimes \bar{\rho}_T))$ is actually a vector bundle over T . This implies that $\tau_{\leq 1} \mathcal{C}^\bullet(\bar{\sigma}^\vee \otimes \bar{\rho}_T)$ is still perfect, and can thus be modeled as a two term complex $[\bar{C}^0 \rightarrow Z^1]$ concentrated in degrees 0, 1, whose terms are finite locally free over T . Note that we have a surjection $Z^1 \twoheadrightarrow H^1(\mathcal{C}^\bullet(\bar{\sigma}^\vee \otimes \bar{\rho}_T)) = \text{Ext}^1(\bar{\sigma}, \bar{\rho}_T)$.

Now let V be the vector bundle over T corresponding to Z^1 . Let $\bar{\rho}_V$ be the pullback of $\bar{\rho}_T$ to V . By analogy with the 2-dimensional case, we want to study a “universal extension” \mathcal{E}_V of $\bar{\rho}_V$ by $\bar{\sigma}$, which should fit in the short exact sequence

$$0 \rightarrow \bar{\rho}_V \rightarrow \mathcal{E}_V \rightarrow \bar{\sigma} \rightarrow 0.$$

To construct this thing, we consider

$$\text{Ext}^1(\bar{\sigma}, \bar{\rho}_T \otimes (Z^1)^\vee) = \text{Ext}^1(\bar{\sigma}, \bar{\rho}_T) \otimes (Z^1)^\vee,$$

which is a quotient of

$$Z^1 \otimes (Z^1)^\vee.$$

This tensor product has a “trace element”, whose image in $\text{Ext}^1(\bar{\sigma}, \bar{\rho}_T \otimes (Z^1)^\vee)$ is an extension class corresponding to a short exact sequence

$$0 \rightarrow \bar{\rho}_T \otimes (Z^1)^\vee \rightarrow \text{extension} \rightarrow \bar{\sigma} \rightarrow 0.$$

If we extend scalars from $\bar{\mathbf{F}}_p$ to \mathcal{O}_V , and then pushforward under the map

$$\rho_T \otimes_{\mathcal{O}_T} (Z^1)^\vee \otimes_{\mathbf{F}_p} \mathcal{O}_V \rightarrow \rho_T \otimes_{\mathcal{O}_T} \mathcal{O}_V$$

induced by the inclusion $(Z^1)^\vee \hookrightarrow \text{Sym}^\bullet(Z^1)^\vee = \mathcal{O}_V$ together with the multiplicative structure of \mathcal{O}_V , we obtain the desired extension

$$0 \rightarrow \bar{\rho}_T \otimes_{\mathcal{O}_T} \mathcal{O}_V \rightarrow \mathcal{E}_V \rightarrow \bar{\sigma} \otimes_{\bar{\mathbf{F}}_p} \mathcal{O}_V \rightarrow 0.$$

Here \mathcal{E}_V is now an étale (φ, Γ) -module over V , and so is classified by a map $V \rightarrow \mathcal{X}_{d, \text{red}}$ (where $d = \dim \bar{\rho}_T + \dim \bar{\sigma}$). The object \mathcal{E}_V will admit various automorphisms, and if we compute them precisely, we can hope to get an embedding of a quotient stack of V into $\mathcal{X}_{d, \text{red}}$. Iterating this construction, we obtain families of étale (φ, Γ) -modules parameterized by quotient stacks of the total spaces of the various vector bundles V appearing via applications of the preceding construction, which will ultimately cover the various $\mathcal{X}_{d, \text{red}}$. (Since any Galois representation over $\bar{\mathbf{F}}_p$ can be written as an iterated extension of irreducible representations.)

7.3. Explicit families and labelling by Serre weights. The iterative construction that we just explained ultimately allows us to get a description of the underlying reduced stacks $\mathcal{X}_{d,\text{red}}$ for general d . Before stating our general result, we return to the case of $K = \mathbf{Q}_p$ and $d = 2$. Again, while the stack \mathcal{X}_2 is really a stack of (φ, Γ) -modules, for the purposes of gaining intuition, it is reasonable to imagine our families of extensions as being extensions of Galois representations, and we will describe things from this perspective without further comment for the rest of this section. For simplicity we assume that $p > 2$ throughout this discussion.

Take $T = \mathbf{G}_m = \text{Spec } \mathbf{F}_p[\alpha, \alpha^{-1}]$ and $\bar{\rho}_T = \text{ur}_\alpha \omega^i$. For the moment let $\bar{\sigma} = 1$ (this will suffice to describe all the geometric phenomena; twisting by a 1-dimensional family $\text{ur}_\beta \omega^j$ for some $0 \leq j < p - 1$ gives the remaining cases). Fix some $0 \leq i < p - 1$. Then

$$\text{Ext}^2(1, \text{ur}_\alpha \omega^i) = H^2(G_{\mathbf{Q}_p}, \text{ur}_\alpha \omega^i)$$

which (for example by Tate local duality) vanishes unless $i = 1$ and $\alpha = 1$, in which case it's 1-dimensional.

We begin with the case that $2 \leq i \leq p - 2$. Then

$$\text{Ext}^2(1, \text{ur}_\alpha \omega^i) = 0,$$

and

$$\text{Ext}^0(1, \text{ur}_\alpha \omega^i) = (\text{ur}_\alpha \omega^i)^{G_{\mathbf{Q}_p}} = 0$$

so actually (by Tate's local Euler characteristic formula) $\text{Ext}^1(1, \text{ur}_\alpha \omega^i)$ has constant rank 1, so $\text{Ext}^1(1, \bar{\rho}_T)$ is a line bundle. Its total space V is then a copy of $\mathbf{A}^1 \times \mathbf{G}_m$ that parameterizes a family of extensions

$$\begin{pmatrix} \text{ur}_\alpha \omega^i & * \\ 0 & 1 \end{pmatrix}.$$

We can add one dimension to this family by forming unramified twists — thus obtaining a family parameterized by $V \times \mathbf{G}_m = \mathbf{A}^1 \times \mathbf{G}_m \times \mathbf{G}_m$, parameterizing extensions of the form

$$\begin{pmatrix} \text{ur}_{\alpha\beta} \omega^i & * \\ 0 & \text{ur}_\beta \end{pmatrix}.$$

There is an action of (a different copy of) $\mathbf{G}_m \times \mathbf{G}_m$ on this family, arising from the action of \mathbf{G}_m by automorphisms on each of the characters $\text{ur}_{\alpha\beta} \omega^i$, and ur_β . If we let (a, b) denote the coordinates on this copy $\mathbf{G}_m \times \mathbf{G}_m$ that is acting, and let x denote the coordinate on \mathbf{A}^1 (thus x provides a coordinate on each of the 1-dimensional spaces $\text{Ext}^1(\text{ur}_\beta, \text{ur}_{\alpha\beta} \omega^i)$) — so that the coordinates on $V = \mathbf{A}^1 \times \mathbf{G}_m \times \mathbf{G}_m$ are (x, α, β) — then this action is given by the formula $(a, b) \cdot (x, \alpha, \beta) = (ab^{-1}x, \alpha, \beta)$. (The fact that the stabilizer of any particular point (x, α, β) is the diagonal copy of \mathbf{G}_m if $x \neq 0$, and the entirety of $\mathbf{G}_m \times \mathbf{G}_m$ when $x = 0$, is a manifestation of the fact that the only automorphisms of a non-split extension of distinct characters are the scalar matrices, while the automorphisms of a split extension (i.e. a direct sum) of distinct characters are the diagonal matrices.) Thus we get an embedding

$$[(\mathbf{A}^1 \times \mathbf{G}_m \times \mathbf{G}_m)/(\mathbf{G}_m \times \mathbf{G}_m)] \hookrightarrow \mathcal{X}_{2,\text{red}},$$

whose source visibly has dimension $3 - 2 = 1$.

We next turn to the case $i = 0$, where Tate local duality gives us that

$$\text{Ext}^2(1, \text{ur}_\alpha) = H^2(G_{\mathbf{Q}_p}, \text{ur}_\alpha) = H^0(G_{\mathbf{Q}_p}, \text{ur}_\alpha \otimes \epsilon^{-1}) = 0$$

for all $\alpha \in \overline{\mathbf{F}}_p^\times$. However, we now run into the issue that Ext^0 is not a vector bundle, because it has rank 0 everywhere except for when $\alpha = 1$, and at this point the rank jumps from 0 to 1; correspondingly, the rank of $\text{Ext}^1(1, \text{ur}_\alpha)$ jumps from 1 to 2. In the framework of Subsection 7.2, this means that we can't choose the complex $C^0 \rightarrow Z^1$ so that $C^0 = 0$ and $Z^1 = \text{Ext}^1$; we have to instead allow C^0 to have rank 1, and Z^1 to have rank 2. If we go through the construction of that subsection, we end up with a family parameterized by the total space V of a rank 2 vector bundle on \mathbf{G}_m . We can add in the unramified twists, to obtain a family parameterized by $V \times \mathbf{G}_m$, and we obtain a map

$$(7.3.1) \quad V \times \mathbf{G}_m \rightarrow \mathcal{X}_{2,\text{red}}$$

classifying the family of Galois representations that $V \times \mathbf{G}_m$ parameterizes. This map won't be an embedding, of course, because various different points on $V \times \mathbf{G}_m$ will give rise to isomorphic Galois representations. Just as in the case $2 \leq i \leq p-2$ considered above, there is a scaling action of $\mathbf{G}_m \times \mathbf{G}_m$ on the various Ext^1 spaces. Also, any two points in of V that lie over the same point $\alpha \in \mathbf{G}_m(\overline{\mathbf{F}}_p)$, and which lie in the same fibre of the surjection

$$(7.3.2) \quad Z_\alpha^1 \rightarrow \text{Ext}^1(1, \text{ur}_\alpha)$$

(here we write Z_α^1 to denote the fibre of Z^1 over the point α) parameterize the same extension of 1 by ur_α . And there are even more isomorphisms that can occur between the representations parameterized by this family: for example, if we restrict to the family of split extensions $\text{ur}_{\alpha\beta} \oplus \text{ur}_\beta$ parameterized by $\mathbf{G}_m \times \mathbf{G}_m$ (the “zero section” of $V \times \mathbf{G}_m$, which we also refer to as the “split locus” of the family), then the points (α, β) and $(\alpha^{-1}, \alpha\beta)$ both parameterize the direct sum $\text{ur}_{\alpha\beta} \oplus \text{ur}_\beta$. Since these “extra isomorphisms” occur only on a proper subvariety of $V \times \mathbf{G}_m$, though, they don't play a role in determining the dimension of the image of (7.3.1). Indeed, since the kernel of the surjection (7.3.2) generically (more precisely, whenever $\alpha \neq 1$) has dimension 1, we find that the image of (7.3.1) is an irreducible constructible substack of $\mathcal{X}_{2,\text{red}}$ of dimension $4 - 2 - 1 = 1$. The structure of this substack along the image of the split locus, as well as along the image locus where $\alpha = 1$, is rather complicated; for example, along the split locus we get a fold singularity (corresponding to the identification of pairs of points (α, β) and $(\alpha^{-1}, \alpha\beta)$), degenerating to some kind of cusp singularities along the points where furthermore $\alpha = 1$. (See [San14] for some computations of the versal rings at these points.)

Finally, in the case $i = 1$, Tate local duality gives

$$\dim \text{Ext}^2(1, \text{ur}_\alpha \omega) = \dim H^2(G_{\mathbf{Q}_p}, \text{ur}_\alpha \omega) = \dim H^0(G_{\mathbf{Q}_p}, \text{ur}_\alpha) = \begin{cases} 1 & \alpha = 1 \\ 0 & \alpha \neq 1 \end{cases}$$

so Ext^2 is no longer a vector bundle. On the other hand, we have

$$H^0(G_{\mathbf{Q}_p}, \text{ur}_\alpha \omega) = 0.$$

Following the strategy of Subsection 7.2, we thus break \mathbf{G}_m up into two pieces: $T_0 := \mathbf{G}_m \setminus \{\alpha = 1\}$, and $T_1 := \{\alpha = 1\}$. Over T_0 , we see that $\text{Ext}^2(1, \text{ur}_\alpha \omega)$ has constant rank 0, and we can construct a family analogously to the case $2 \leq i \leq p-1$, parameterized by the total space of a line bundle over $\mathbf{G}_m \setminus \{\alpha = 1\}$. Adding in the unramified twists, this leads to an embedding

$$[\mathbf{A}^1 \times (\mathbf{G}_m \setminus \{\alpha = 1\}) \times \mathbf{G}_m / (\mathbf{G}_m \times \mathbf{G}_m)] \hookrightarrow \mathcal{X}_{2,\text{red}}.$$

Passing now to T_1 (which is just a single point! — over which $\mathrm{Ext}^2(1, \mathrm{ur}_\alpha \omega)$ has constant rank 1), we obtain a family parameterized by the 2-dimensional vector space $\mathrm{Ext}^1(1, \omega)$. Again adding in unramified twists, we obtain an embedding

$$[(\mathbf{A}^2 \times \mathbf{G}_m 0 / (\mathbf{G}_m \times \mathbf{G}_m))] \hookrightarrow \mathcal{X}_{2, \mathrm{red}}.$$

Each of the finite number of families that we have just described (along with the families obtained by taking a twist of one of these families by ω^j for some $j = 1, \dots, p-2$) gives rise to a 1-dimensional irreducible constructible substack of $\mathcal{X}_{2, \mathrm{red}}$. By construction, every $\overline{\mathbf{F}}_p$ -point of $\mathcal{X}_{2, \mathrm{red}}$ that corresponds to a *reducible* Galois representation lies exactly one of these subsets.

We have already seen that the $\overline{\mathbf{F}}_p$ -points corresponding to *irreducible* Galois representations lie in a finite union of images of the 0-dimensional stack $[\mathbf{G}_m / \mathbf{G}_m]$. We thus deduce that $\mathcal{X}_{2, \mathrm{red}}$ is 1-dimensional, and that the closure of each of the 1-dimensional and mutually disjoint constructible substacks that we've just constructed must be an irreducible component of $\mathcal{X}_{2, \mathrm{red}}$. We now explain how we can label these components by so-called *Serre weights*.

Temporarily return to the case of general K , d , and let k be the residue field of K . Then by definition, a Serre weight is an (isomorphism class of) irreducible $\overline{\mathbf{F}}_p$ -representations of $\mathrm{GL}_n(k)$. These are determined by their highest weights, which are tuples of integers $\{k_{\overline{\sigma}, i}\}_{\overline{\sigma}: k \hookrightarrow \overline{\mathbf{F}}_p, 1 \leq i \leq d}$ with the properties that

- $p-1 \geq k_{\overline{\sigma}, i} - k_{\overline{\sigma}, i+1} \geq 0$ for each $1 \leq i \leq d-1$, and
- $p-1 \geq k_{\overline{\sigma}, d} \geq 0$, and not every $k_{\overline{\sigma}, d}$ is equal to $p-1$.

We refer the reader to Lecture 9 for some explanations and motivation for the appearance of the representation theory of $\mathrm{GL}_n(k)$ in the geometry of our stacks.

Returning to the case that $K = \mathbf{Q}_p$ and $d = 2$, the Serre weights are the representations $\det^j \mathrm{Sym}^{i-1} \overline{\mathbf{F}}_p^2$ with $0 \leq j < p-1$, $1 \leq i \leq p$; in the notation above, these have highest weight $(k_1, k_2) = (j+i-1, j)$. Our labelling of components by Serre weights is then as follows.⁴ An irreducible component which generically parameterizes representations of the form

$$\begin{pmatrix} \mathrm{ur}_{\alpha\beta} \omega^{i+j} & * \\ 0 & \mathrm{ur}_{\beta} \omega^j \end{pmatrix}$$

is labelled by the highest weight $(i+j-1, j)$ (that is, by the representation $\det^j \mathrm{Sym}^{i-1} \overline{\mathbf{F}}_p^2$). This is unambiguous except for the possibilities $i = 1, p$, which we distinguish as follows. In the discussion above, we saw that if $i = 1$ there is one irreducible component for which $\alpha = 1$ at a dense set of points, and we label this component by $\det^j \mathrm{Sym}^{p-1} \overline{\mathbf{F}}_p^2$; and there is one component where $\alpha \neq 1$ at a dense set of points, which we label by $\det^j \mathrm{Sym}^0 \overline{\mathbf{F}}_p^2$.

We can inductively extend the the construction technique that we have just described for $d = 2$, so as to prove the following general theorem. The notion of being “generically maximally nonsplit of niveau one” means that for an open set of

⁴For accuracy, in case the reader tries to carefully compare our present discussion with the corresponding discussion in [EG19], we note that our labelling here differs from the one given in that reference; the component that here we label by the Serre weight σ will there be labelled by the Serre weight $\sigma^\vee \otimes \det^{-1}$. We have chosen our present convention just because it is notationally easier to work with extensions of 1 by powers of ω rather than with extensions of powers of ω^{-1} by 1. If you like, we have applied the “Cartier duality” involution — given by $\overline{\rho} \mapsto \overline{\rho}^\vee \otimes \omega$ — to $\mathcal{X}_{2, \mathrm{red}}$.

points, the corresponding G_K -representation is a successive extension of characters in a unique way; to such a representation one can associate a Serre weight \underline{k} , via a straightforward extension of the labelling we just explained in the case $K = \mathbf{Q}_p$ and $d = 2$.

7.3.3. Theorem.

- (1) The Ind-algebraic stack $\mathcal{X}_{d,\text{red}}$ is an algebraic stack, of finite presentation over \mathbf{F} .
- (2) We can write $(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p}$ as a union of closed algebraic substacks of finite presentation over $\overline{\mathbf{F}}_p$

$$(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p} = \mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\text{small}} \cup \bigcup_{\underline{k}} \mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\underline{k}},$$

where:

- $\mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\text{small}}$ is empty if $d = 1$, and otherwise is non-empty of dimension strictly less than $[K : \mathbf{Q}_p]d(d-1)/2$.
 - each $\mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\underline{k}}$ is a closed irreducible substack of dimension $[K : \mathbf{Q}_p]d(d-1)/2$, and is generically maximally nonsplit of niveau one and weight \underline{k} .
- (3) If we fix an irreducible representation $\overline{\alpha} : G_K \rightarrow \text{GL}_a(\overline{\mathbf{F}}_p)$ (for some $a \geq 1$), then the locus of $\overline{\rho}$ in $\mathcal{X}_{d,\text{red}}(\overline{\mathbf{F}}_p)$ for which $\dim \text{Ext}_{G_K}^2(\overline{\alpha}, \overline{\rho}) \geq r$ is of dimension at most

$$[K : \mathbf{Q}_p]d(d-1)/2 - r.$$

7.3.4. *Remark.* Note that in part (3), the locus of points in question corresponds to a closed substack of $(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p}$, by upper-semicontinuity of fibre dimension.

7.3.5. *Remark.* Since the $\mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\underline{k}}$ are irreducible, have dimension equal to that of $(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p}$, and have pairwise disjoint open substacks (corresponding to maximally nonsplit representations of niveau 1 and weight \underline{k}), they are in fact distinct irreducible components of $(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p}$.

7.3.6. *Remark.* We can, and will, be much more precise about the structure of $\mathcal{X}_{d,\text{red}}$. Namely, in Lecture 9 we combine Theorem 7.3.3 with the results of Lecture 8 to show that $\mathcal{X}_{d,\text{red}}$ is equidimensional of dimension $[K : \mathbf{Q}_p]d(d-1)/2$; accordingly, the irreducible components of $(\mathcal{X}_{d,\text{red}})_{\overline{\mathbf{F}}_p}$ are precisely the $\mathcal{X}_{d,\text{red},\overline{\mathbf{F}}_p}^{\underline{k}}$, and in particular are in bijection with the Serre weights \underline{k} .

7.3.7. *Remark.* The upper bound of Theorem 7.3.3 (3) is quite crude when $[K : \mathbf{Q}_p] > 1$, although it is reasonably sharp in the case $K = \mathbf{Q}_p$. However, it suffices for our purposes. (In fact, in the inductive proof of Theorem 7.3.3 we use a slightly more complicated upper bound to control the dimensions of various extension groups, but the stated bound is all that we will need going forward.)

LECTURE 8. CRYSTALLINE LIFTS

8.1. **Crystalline lifts.** The following theorem is the main result of this lecture.

8.1.1. **Theorem.** *If $\overline{\rho} : G_K \rightarrow \text{GL}_d(\overline{\mathbf{F}}_p)$ is continuous, then there exists $\rho^\circ : G_K \rightarrow \text{GL}_d(\overline{\mathbf{Z}}_p)$ lifting $\overline{\rho}$ such that the corresponding p -adic Galois representation $\rho : G_K \rightarrow \text{GL}_d(\overline{\mathbf{Q}}_p)$ is crystalline with regular Hodge-Tate weights.*

8.1.2. *Remark.* In fact, the lift ρ° in Theorem 8.1.1 can be chosen to be potentially diagonalizable in the sense of [BLGGT14], which implies in turn that if $p \nmid 2d$, then $\bar{\rho}$ can be globalized to come from an automorphic form (see [EG19, Thm. 1.2.3]).

While the statement of Theorem 8.1.1 is purely local (in the sense that the representation $\bar{\rho}$ is fixed), and our proof is for the most part via a local argument, we will also make crucial use of Theorem 7.3.3.

We begin by explaining the “obvious” strategy to prove Theorem 8.1.1. Firstly consider the case that $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\bar{\mathbf{F}}_p)$ is irreducible. Then it is easy to show that $\bar{\rho}$ is induced from a character of the unramified extension K'/K of degree d , and furthermore it is easy to show that this character can be lifted to a crystalline character, and because K'/K is unramified, the induction of a crystalline character of $G_{K'}$ to G_K gives a crystalline representation. Furthermore, one has considerable control over the Hodge–Tate weights of such a lift.

An immediate consequence is that all semisimple $\bar{\rho}$ have crystalline lifts. However, not all mod p representations are semisimple, so the remaining problem is to show that we can lift extensions of representations. To get a feeling for this, consider the two-dimensional case, and so suppose that $\bar{\rho} : G_K \rightarrow \mathrm{GL}_2(\bar{\mathbf{F}}_p)$ is of the form

$$\begin{pmatrix} \bar{\chi}_1 & * \\ 0 & \bar{\chi}_2 \end{pmatrix}.$$

Choose, as we may, crystalline lifts $\chi_1 : G_K \rightarrow \bar{\mathbf{Z}}_p^\times$, $\chi_2 : G_K \rightarrow \bar{\mathbf{Z}}_p^\times$ of $\bar{\chi}_1$, $\bar{\chi}_2$ respectively, and suppose that the labelled Hodge–Tate weights of χ_1 are *slightly greater* than those of χ_2 , in the sense that for each embedding $K \hookrightarrow \bar{\mathbf{Q}}_p$ the corresponding Hodge–Tate weight of χ_1 is greater than that of χ_2 , and for at least one embedding, the gap is greater than 1. Then any extension of χ_2 by χ_1 is automatically crystalline (for example by the formulae in [Nek93, Prop. 1.24]). Since extensions are classified by H^1 , it is therefore enough to show that given any class in $H^1(G_K, \bar{\chi}_1 \bar{\chi}_2^{-1})$, we can choose χ_1 , χ_2 in such a way that this class can be lifted to $H^1(G_K, \chi_1 \chi_2^{-1})$.

As we will soon see, this can always be arranged, and in “most” cases the natural map $H^1(G_K, \chi_1 \chi_2^{-1}) \rightarrow H^1(G_K, \bar{\chi}_1 \bar{\chi}_2^{-1})$ is actually surjective. It is then natural to imagine that one could inductively prove the existence of lifts in all dimensions by induction on the number of Jordan–Hölder factors of $\bar{\rho}$. Quite a few people have tried to do this, but no-one succeeded, as far as we know. The basic problem seems to be that it is in general not possible to fix the lifts of each Jordan–Hölder factor when arguing inductively; rather, one keeps having to go back and adjust the previous choices in light of the next step.

For concreteness, here are some examples (due to Tong Liu) which are difficult to do by hand, and give a sense of why it’s hard to work inductively. For the rest of the lecture we write ε for the cyclotomic character, and $\bar{\varepsilon}$ for the mod p cyclotomic character. Firstly, consider

$$(8.1.3) \quad \bar{W} = \begin{pmatrix} \bar{\varepsilon} & * & * & * \\ & \bar{\varepsilon} & 0 & * \\ & & \bar{\varepsilon} & * \\ & & & 1 \end{pmatrix}$$

where $*$ could be zero or not. It is quite easy to convince yourself that there are enough “degrees of freedom” (by choosing for example different unramified twists of the characters that you lift to, and lifting to different extension classes) to manage to make the lift, but quite tricky to come up with an actual argument. If that one is too easy, you can try something like

$$\begin{pmatrix} \bar{\varepsilon}^2 & * & * & * \\ & \bar{\varepsilon}^2 & 0 & * \\ & & \bar{\varepsilon}^2 & * \\ & & & \overline{W} \end{pmatrix}$$

The key to our approach is the following theorem. Note that when we use this theorem inductively, we are not guaranteeing that you can fix lifts of each Jordan–Hölder factor once and for all before considering the various possible extension classes, and this increased flexibility is important in our argument.

8.1.4. Theorem. *Suppose given a representation $\bar{\rho}_d : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{F}}_p)$ that admits a lift $\rho_d^\circ : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{Z}}_p)$ which is crystalline with labelled Hodge–Tate weights $\underline{\lambda}$. Let $0 \rightarrow \bar{\rho}_d \rightarrow \bar{\rho}_{d+a} \rightarrow \bar{\alpha} \rightarrow 0$ be any extension of G_K -representations over $\overline{\mathbf{F}}_p$, with $\bar{\alpha} : G_K \rightarrow \mathrm{GL}_a(\overline{\mathbf{F}}_p)$ irreducible, and let $\alpha^\circ : G_K \rightarrow \mathrm{GL}_a(\overline{\mathbf{Z}}_p)$ be any crystalline lifting of $\bar{\alpha}$ with labelled Hodge–Tate weights $\underline{\lambda}'$, which we assume to be slightly less than $\underline{\lambda}$.*

Then we may find a lifting of the given extension to an extension

$$0 \rightarrow \theta_d^\circ \rightarrow \theta_{d+a}^\circ \rightarrow \alpha^\circ \rightarrow 0$$

of G_K -representations over $\overline{\mathbf{Z}}_p$, where $\theta_d^\circ : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{Z}}_p)$ again has the property that the associated p -adic representation $\theta_d : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{Q}}_p)$ is crystalline with labelled Hodge–Tate weights $\underline{\lambda}$. Furthermore, θ_{d+a}° is crystalline, and we may choose θ_d° to lie on the same irreducible component of $\mathrm{Spec} R_{p_d}^{\underline{\lambda}, \mathrm{crys}}$ that ρ_d° does.

Given Theorem 8.1.4, it is easy to prove Theorem 8.1.1 by induction on d . In fact, without much additional difficulty one can prove the following stronger version, which is useful in applications.

8.1.5. Theorem. *Let K/\mathbf{Q}_p be a finite extension, and let $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{F}}_p)$ be a continuous representation. Then $\bar{\rho}$ admits a lift to a crystalline representation $\rho^\circ : G_K \rightarrow \mathrm{GL}_d(\overline{\mathbf{Z}}_p)$ of some regular labelled Hodge–Tate weights $\underline{\lambda}$. Furthermore:*

- (1) ρ° can be taken to be potentially diagonalizable.
- (2) If every Jordan–Hölder factor of $\bar{\rho}$ is one-dimensional, then ρ° can be taken to be ordinary.
- (3) ρ° can be taken to be potentially diagonalizable, and $\underline{\lambda}$ can be taken to be a lift of a Serre weight.
- (4) ρ° can be taken to be potentially diagonalizable, and $\underline{\lambda}$ can be taken to have arbitrarily spread-out Hodge–Tate weights: that is, for any $C > 0$, we can choose ρ° such that for each $\sigma : K \hookrightarrow \overline{\mathbf{Q}}_p$, we have $\lambda_{\sigma, i} - \lambda_{\sigma, i+1} \geq C$ for each $1 \leq i \leq n - 1$.

We now discuss the proof of Theorem 8.1.4. Return to the two-dimensional case, and write $\bar{\chi} = \bar{\chi}_1 \bar{\chi}_2^{-1}$, $\chi = \chi_1 \chi_2^{-1}$, so that we want to lift classes in $H^1(G_K, \bar{\chi})$ to $H^1(G_K, \chi)$. Fix a sufficiently large finite extension E/\mathbf{Q}_p with ring of integers \mathcal{O} ,

uniformizer ϖ , and residue field \mathbf{F} , so that in particular χ takes values in \mathcal{O}^\times . Taking cohomology of

$$0 \rightarrow \mathcal{O}(\chi) \xrightarrow{\varpi} \mathcal{O}(\chi) \rightarrow \mathbf{F}(\bar{\chi}) \rightarrow 0,$$

we see that we have an exact sequence

$$H^1(G_K, \chi) \otimes_{\mathcal{O}} \mathbf{F} \rightarrow H^1(G_K, \bar{\chi}) \rightarrow H^2(G_K, \chi)[\varpi].$$

By Tate local duality, $H^2(G_K, \chi)$ is dual to $H^0(G_K, (E/\mathcal{O})(\chi^{-1})(1))$, with the (1) denoting a Tate twist, so we see in particular that if $\bar{\chi} \neq \bar{\varepsilon}$, then there is nothing to prove. However, if $\bar{\chi} = \bar{\varepsilon}$, then there *is* some work to do.

One thing that we could do in this case is to take $\chi = \varepsilon$, in which case $H^2(G_K, \chi)[\varpi]$ vanishes, and we see that we can lift to semistable (not necessarily crystalline) extensions. One might wonder if this means that in general we should just relax things and try to only produce semistable lifts, rather than crystalline ones, but in practice this doesn't seem to be the case - the inductive arguments still run into trouble on examples like (8.1.3) above. Indeed, let us try to handle (8.1.3) by lifting inductively, allowing ourselves to use semistable lifts. The bottom right 3×3 matrix is a sum of two extensions with $\bar{\chi} = \bar{\varepsilon}$, so we've just seen that we can lift it to a representation U of the form

$$U = \begin{pmatrix} \varepsilon & 0 & * \\ & \varepsilon & * \\ & & 1 \end{pmatrix}.$$

Then

$$\bar{W} = \begin{pmatrix} \bar{\varepsilon} & * \\ & \bar{U} \end{pmatrix},$$

so we could try to consider lifts of the form

$$W = \begin{pmatrix} \varepsilon & * \\ & U \end{pmatrix}.$$

However, if the classes giving the two extensions of $\bar{\varepsilon}$ by $\bar{\varepsilon}$ aren't twists by $\bar{\varepsilon}$ of the unramified extension of the trivial character by itself, then these lifts won't be semistable (because the only semistable extensions of the trivial character by itself are unramified). So we are forced instead to consider something like

$$W = \begin{pmatrix} \varepsilon^p & * \\ & U \end{pmatrix},$$

and then we run into difficulties with the existence of torsion in H^2 .

So let's return to the case of $\bar{\chi} = \bar{\varepsilon}$, and consider the problem of finding crystalline lifts. Rather than taking $\chi = \varepsilon$, we take χ to be an unramified twist of ε^p , so that all extensions are crystalline. In this case one finds that $H^2(G_K, \chi)[\varpi]$ is 1-dimensional, so there is a possible obstruction to lifting; and indeed $H^1(G_K, \bar{\varepsilon})$ is 2-dimensional, while $H^1(G_K, \chi) \otimes_{\mathcal{O}} \mathbf{F}$ is 1-dimensional, so no fixed χ can provide all the lifts that we need. It is possible to analyse the situation directly using Kummer theory, and one finds that given any class in $H^1(G_K, \bar{\varepsilon})$, one can choose χ so that a lift exists.

It is then natural to consider the moduli space of all such χ , and the universal extension over this space. The universal χ is the character $\chi_t : G_K \rightarrow \mathcal{O}[[t]]^\times$ given by $\chi_t = \varepsilon^p \lambda_{1+t}$, where λ_x is the unramified character sending a (geometric) Frobenius element to x . Then by Tate local duality and local class field theory, it

is easy to see that $H^2(G_K, \chi_t) = \mathcal{O}[[t]]/(p, t) = \mathcal{O}/p$. One checks that $H^1(G_K, \chi_t)$ is a free \mathcal{O} -module of rank 1, but that the corresponding “universal family” of extensions over $\mathrm{Spf} \mathcal{O}[[t]]$ is not in fact universal; more precisely, the fibre over any given $x : \mathcal{O}[[t]] \rightarrow \mathcal{O}$, $t \mapsto a$ sits in a short exact sequence

$$(8.1.6) \quad 0 \rightarrow H^1(G_K, \chi_t) \otimes_{\mathcal{O}[[t]], x} \mathcal{O} \rightarrow H^1(G_K, \varepsilon^p \lambda_{1+a}) \rightarrow \mathcal{O}[[t]]/(p, t) \rightarrow 0.$$

So we see would like to have a way of “rescaling” $H^1(G_K, \chi_t)$ by dividing by (p, t) , so as to have a universal family that sees every extension; and then we could try to check that this family surjects onto $H^1(G_K, \bar{\varepsilon})$. To do this, we blow up $\mathrm{Spf} \mathcal{O}[[t]]$ along (p, t) , giving a p -adic formal scheme \mathcal{X} , which is flat over \mathbf{Z}_p , and for which (p, t) pulls back to an ideal sheaf \mathcal{I} . Over \mathcal{X} , we do have a universal family $H^1(G_K, \chi_{\mathcal{X}})$, and one can check that its base change to any point $\tilde{x} : \mathrm{Spf} \mathcal{O} \rightarrow \mathcal{X}$ gives the full H^1 .

More explicitly, since H^0 plays no role, we may assume that the Herr complex is computed by a complex $C^1 \rightarrow C^2$ of locally free modules over $\mathcal{O}[[t]]$. Since $H^2 = \mathcal{O}[[t]]/(p, t)$, we can imagine that this complex looks like

$$\mathcal{O}[[t]]^2 \xrightarrow{(p, t)} \mathcal{O}[[t]].$$

(Note that this is consistent with (8.1.6).) In particular, the image of the boundary map is the ideal $I = (p, t)$. When we blow up $\mathrm{Spec} \mathcal{O}[[t]]$ at (p, t) to obtain a scheme X , this complex pulls back to a complex

$$\tilde{C}^1 = \mathcal{O}_X^{\oplus 2} \rightarrow \mathcal{O}_X = \tilde{C}^2,$$

and the image of the boundary map is now locally free, equal to the ideal sheaf \mathcal{I} of the exceptional divisor. Thus the formation of H^1 of this pulled-back complex commutes with arbitrary base change (the surjection $\tilde{C}^1 \rightarrow \mathcal{I}$ induced by the boundary map is split, since the target is locally free).

Given a class in $H^1(G_K, \bar{\varepsilon})$ arising from a cocycle \bar{c} , we lift \bar{c} to a map $c : \mathcal{O}[[t]] \rightarrow C^1$ which specializes at $(p, t) = 0$ to a cocycle. This lifts to a map $\tilde{c} : \mathcal{O}_X \rightarrow \tilde{C}^1$, which composed with the boundary gives a map $\delta \tilde{c} : \mathcal{O}_X \rightarrow \mathcal{I}$. Since \mathcal{I} is not the trivial line bundle, this map (thought of as a section of the line bundle associated to \mathcal{I}) has a zero at some point, which means it gives a class in H^1 of the base-change of \tilde{C}^\bullet to this point, which then (by the base-change observation of the preceding paragraph) lifts to a class in H^1 of \tilde{C}^\bullet . (Basically we see that $\tilde{C}^1 = \tilde{Z}^1 \oplus \mathcal{I}$, and we replace \tilde{c} by its projection to the first component, so that it becomes a cocycle; this doesn’t change its value at the point where it vanishes, and thus means that this cocycle lifts the cocycle \bar{c} . This projection achieves the “division by (p, t) ”.)

The proof of Theorem 8.1.4 is via a generalisation of this construction, with the following theorem being the general analogue that we need of the fact used above that $H^2(G_K, \chi_t) = \mathcal{O}[[t]]/(p, t)$. (To see the relationship with the previous argument, which used non-triviality of the line bundle associated to \mathcal{I} , note that if $\delta \tilde{c}$ induced a surjection $\mathcal{O}_X \rightarrow \mathcal{I}$, then we would see that c induces a surjection $\mathcal{O}[[t]] \rightarrow I$, hence I would be principal, and $H^2 = \mathcal{O}[[t]]/I$ would be supported in codimension one, whereas in fact it is supported only in codimension two.)

8.1.7. Theorem. *For any regular tuple of labelled Hodge–Tate weights $\underline{\lambda}$, and any fixed irreducible representation $\bar{\alpha} : G_K \rightarrow \mathrm{GL}_a(\bar{\mathbf{F}}_p)$, the locus of points $x \in$*

Spec $R^{\lambda, \text{crys}}/\varpi$ for which

$$\dim_{\kappa(x)} \kappa(x) \otimes_{R_{\bar{\rho}}} \text{Ext}_{G_K}^2(\bar{\alpha}, \rho^{\text{univ}}) \geq r$$

has codimension at least r .

Proof. See [EG19, Thm. 6.1.1]. The basic idea is that since the stack $\mathcal{X}_d^{\lambda, \text{crys}}$ is a p -adic formal algebraic stack, its special fibre is an algebraic stack, and the ring $R^{\lambda, \text{crys}}/\varpi$ is an effective versal ring at some finite type point. It therefore suffices to prove a corresponding codimension bound on the support of $\text{Ext}_{G_K}^2(\bar{\alpha}, -)$ in the special fibre of \mathcal{X}_d , and this is provided by Theorem 7.3.3 (3). \square

LECTURE 9. GEOMETRIC BREUIL–MÉZARD

As in the previous lectures, we will concentrate on the potentially crystalline case when formulating the Breuil–Mézard conjecture; most of what we say goes over unchanged to the potentially semistable version, and we refer the reader to [EG19, §8] for the details.

9.1. The irreducible components of $\mathcal{X}_{d, \text{red}}$. We can now complete the analysis of the irreducible components of $\mathcal{X}_{d, \text{red}}$. Recall that by Theorem 7.3.3, $\mathcal{X}_{d, \text{red}}$ is an algebraic stack of finite presentation over \mathbf{F} , and has dimension $[K : \mathbf{Q}_p]d(d-1)/2$. Furthermore, for each Serre weight \underline{k} , there is a corresponding irreducible component $\mathcal{X}_{d, \text{red}, \bar{\mathbf{F}}_p}^{\underline{k}}$ of $(\mathcal{X}_{d, \text{red}})_{\bar{\mathbf{F}}_p}$, and the components for different weights \underline{k} are distinct. Using our results on crystalline lifts, we can now show that these are the only irreducible components. The following is [EG19, Thm. 6.5.1].

9.1.1. Theorem. *$\mathcal{X}_{d, \text{red}}$ is equidimensional of dimension $[K : \mathbf{Q}_p]d(d-1)/2$, and the irreducible components of $(\mathcal{X}_{d, \text{red}})_{\bar{\mathbf{F}}_p}$ are precisely the various closed substacks $\mathcal{X}_{d, \text{red}, \bar{\mathbf{F}}_p}^{\underline{k}}$; in particular, $(\mathcal{X}_{d, \text{red}})_{\bar{\mathbf{F}}_p}$ is maximally nonsplit of niveau 1. Furthermore each $\mathcal{X}_{d, \text{red}, \bar{\mathbf{F}}_p}^{\underline{k}}$ can be defined over \mathbf{F} , i.e. is the base change of an irreducible component $\mathcal{X}_{d, \text{red}}^{\underline{k}}$ of $\mathcal{X}_{d, \text{red}}$.*

Proof. To see that the $\mathcal{X}_{d, \text{red}, \bar{\mathbf{F}}_p}^{\underline{k}}$ may all be defined over \mathbf{F} , we need to show that the action of $\text{Gal}(\bar{\mathbf{F}}_p/\mathbf{F})$ on the irreducible components of $(\mathcal{X}_{d, \text{red}})_{\bar{\mathbf{F}}_p}$ is trivial. This follows immediately by considering its action on the maximally nonsplit representations of niveau 1 (since the action of $\text{Gal}(\bar{\mathbf{F}}_p/\mathbf{F})$ preserves the property of being maximally nonsplit of niveau 1 and weight \underline{k}). It is therefore enough to prove that each irreducible component of $(\mathcal{X}_{d, \text{red}})_{\bar{\mathbf{F}}_p}$ is of dimension of at least $[K : \mathbf{Q}_p]d(d-1)/2$; this follows by choosing a closed point not contained in any other irreducible component, and noting that (by Theorem 8.1.1) it is contained in the special fibre of some $\mathcal{X}_d^{\text{crys}, \underline{\lambda}}$ with $\underline{\lambda}$ regular. \square

We can also prove [EG19, Prop. 6.5.2]:

9.1.2. Proposition. *\mathcal{X}_d is not a p -adic formal algebraic stack.*

Proof. Assume that \mathcal{X}_d is a p -adic formal algebraic stack, so that its special fibre $\bar{\mathcal{X}}_d := \mathcal{X}_d \times_{\mathcal{O}} \mathbf{F}$ is an algebraic stack, which is furthermore of finite type over \mathbf{F} . Since the underlying reduced substack of $\bar{\mathcal{X}}_d$ is $\mathcal{X}_{d, \text{red}}$, which is equidimensional of dimension $[K : \mathbf{Q}_p]d(d-1)/2$, we see that $\bar{\mathcal{X}}_d$ also has dimension $[K : \mathbf{Q}_p]d(d-1)/2$.

Computing with versal rings, we see that for every $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})$ the unrestricted framed deformation ring $R_{\bar{\rho}}^{\square}/\varpi$ must have dimension $d^2 + [K : \mathbf{Q}_p]d(d-1)/2$. However, it is known that there are representations $\bar{\rho}$ for which $R_{\bar{\rho}}^{\square}/\varpi$ is formally smooth of dimension $d^2 + [K : \mathbf{Q}_p]d^2$ (see for example [All19, Lem. 3.3.1]). Thus we must have $d^2 = d(d-1)/2$, a contradiction. \square

9.2. The qualitative geometric Breuil–Mézard conjecture. If λ is a regular Hodge type, and τ is any inertial type, then the stack $\mathcal{X}_d^{\mathrm{crys}, \lambda, \tau}$ is a finite type p -adic formal algebraic stack over \mathcal{O} , which is \mathcal{O} -flat and equidimensional of dimension $1 + [K : \mathbf{Q}_p]d(d-1)/2$. It follows that its special fibre $\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$ is an algebraic stack over \mathbf{F} which is equidimensional of dimension $[K : \mathbf{Q}_p]d(d-1)/2$. Since $\mathcal{X}_d^{\mathrm{crys}, \lambda, \tau}$ is a closed substack of \mathcal{X}_d , its special fibre $\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$ is a closed substack of the special fibre $\bar{\mathcal{X}}_d$, and its irreducible components (with the induced reduced substack structure) are therefore closed substacks of the algebraic stack $\bar{\mathcal{X}}_{d, \mathrm{red}}$.

Since $\bar{\mathcal{X}}_{d, \mathrm{red}}$ is equidimensional of dimension $[K : \mathbf{Q}_p]d(d-1)/2$, it follows that the irreducible components of $\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$ are irreducible components of $\bar{\mathcal{X}}_{d, \mathrm{red}}$, and are therefore of the form $\bar{\mathcal{X}}_{d, \mathrm{red}}^{\underline{k}}$ for some Serre weight \underline{k} .

For each \underline{k} , we write $\mu_{\underline{k}}(\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau})$ for the multiplicity of $\bar{\mathcal{X}}_{d, \mathrm{red}}^{\underline{k}}$ as a component of $\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$. We write $Z_{\mathrm{crys}, \lambda, \tau} = Z(\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau})$ for the corresponding cycle, i.e. for the formal sum

$$(9.2.1) \quad Z_{\mathrm{crys}, \lambda, \tau} = \sum_{\underline{k}} \mu_{\underline{k}}(\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}) \cdot \bar{\mathcal{X}}_d^{\underline{k}},$$

which we regard as an element of the finitely generated free abelian group $\mathbf{Z}[\mathcal{X}_{d, \mathrm{red}}]$ whose generators are the irreducible components $\bar{\mathcal{X}}_d^{\underline{k}}$.

Fix some representation $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})$, corresponding to a point $x : \mathrm{Spec} \mathbf{F} \rightarrow \mathcal{X}_d$. For each regular Hodge type λ and inertial type τ , we have an effective versal morphism $\mathrm{Spec} R_{\bar{\rho}}^{\mathrm{crys}, \lambda, \tau}/\varpi \rightarrow \bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$. For each \underline{k} we set

$$\mathcal{C}_{\underline{k}}(\bar{\rho}) := \mathrm{Spf} R_{\bar{\rho}}^{\square} \times_{\mathcal{X}_d} \mathcal{X}_d^{\underline{k}},$$

which we regard as a cycle of dimension $d^2 + [K : \mathbf{Q}_p]d(d-1)/2$ in $\mathrm{Spec} R_{\bar{\rho}}/\varpi$ (note that since $\mathcal{X}_d^{\underline{k}}$ is algebraic, it has effective versal rings, so we really get a subscheme of $\mathrm{Spec} R_{\bar{\rho}}/\varpi$, rather than of $\mathrm{Spf} R_{\bar{\rho}}/\varpi$). The following theorem gives a qualitative version of the refined Breuil–Mézard conjecture [EG14, Conj. 4.2.1]. While its statement is purely local, we do not know how to prove it without making use of the stack \mathcal{X}_d .

9.2.2. Theorem. *Let $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F})$ be a continuous representation. Then there are finitely many cycles of dimension $d^2 + [K : \mathbf{Q}_p]d(d-1)/2$ in $\mathrm{Spec} R_{\bar{\rho}}^{\square}/\varpi$ such that for any regular Hodge type λ and any inertial type τ , each of the special fibres $\mathrm{Spec} R_{\bar{\rho}}^{\mathrm{crys}, \lambda, \tau}/\varpi$ is set-theoretically supported on some union of these cycles.*

Proof. We have $\mathrm{Spec} R_{\bar{\rho}}^{\mathrm{crys}, \lambda, \tau}/\varpi = \mathrm{Spf} R_{\bar{\rho}}^{\square} \times_{\mathcal{X}_d} \bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}$. It follows from (9.2.1), together with the definition of $\mathcal{C}_{\underline{k}}(\bar{\rho})$, that we may write the underlying cycle as

$$(9.2.3) \quad Z(\mathrm{Spec} R_{\bar{\rho}}^{\mathrm{crys}, \lambda, \tau}/\varpi) = \sum_{\underline{k}} \mu_{\underline{k}}(\bar{\mathcal{X}}_d^{\mathrm{crys}, \lambda, \tau}) \cdot \mathcal{C}_{\underline{k}}(\bar{\rho}).$$

The theorem follows immediately (taking our finite set of cycles to be the $\mathcal{C}_{\underline{k}}(\bar{\rho})$). \square

We can regard this theorem as isolating the “refined” part of [EG14, Conj. 4.2.1]; that is, we have taken the original numerical Breuil–Mézard conjecture, formulated a geometric refinement of it, and then removed the numerical part of the conjecture. The numerical part of the conjecture consists of relating the multiplicities $\mu_{\underline{k}}(\overline{\mathcal{X}}_d^{\text{crys}, \lambda, \tau})$ to the representation theory of $\text{GL}_n(k)$, as we now recall.

The following theorem is essentially due to Schneider–Zink [SZ99].

9.2.4. Theorem. *Let $\tau : I_K \rightarrow \text{GL}_d(\overline{\mathbf{Q}}_p)$ be an inertial type. Then there is a finite-dimensional smooth irreducible $\overline{\mathbf{Q}}_p$ -representation $\sigma^{\text{crys}}(\tau)$ with the property that if π is an irreducible smooth $\overline{\mathbf{Q}}_p$ -representation of $\text{GL}_d(K)$, then the $\overline{\mathbf{Q}}_p$ -vector space $\text{Hom}_{\text{GL}_d(\mathcal{O}_K)}(\sigma^{\text{crys}}(\tau), \pi)$ has dimension at most 1, and is nonzero precisely if $\text{rec}_p(\pi)|_{I_F} \cong \tau$, and $N = 0$ on $\text{rec}_p(\pi)$.*

For each regular Hodge type $\underline{\lambda}$ we let $W(\underline{\lambda})$ be the corresponding representation of $\text{GL}_d(\mathcal{O}_K)$, defined as follows: For each $\sigma : K \hookrightarrow \overline{\mathbf{Q}}_p$, we write $\xi_{\sigma, i} = \lambda_{\sigma, i} - (d - i)$, so that $\xi_{\sigma, 1} \geq \dots \geq \xi_{\sigma, d}$. We view each $\xi_{\sigma} := (\xi_{\sigma, 1}, \dots, \xi_{\sigma, d})$ as a dominant weight of the algebraic group GL_d (with respect to the upper triangular Borel subgroup), and we write $M_{\xi_{\sigma}}$ for the algebraic \mathcal{O}_K -representation of $\text{GL}_d(\mathcal{O}_K)$ of highest weight ξ_{σ} . Then we define $L_{\underline{\lambda}} := \bigotimes_{\sigma} M_{\xi_{\sigma}} \otimes_{\mathcal{O}_K, \sigma} \mathcal{O}$.

For each τ we let $\sigma^{\text{crys}, \circ}(\tau)$ denote a choice of $\text{GL}_d(\mathcal{O}_K)$ -stable \mathcal{O} -lattice in $\sigma^{\text{crys}}(\tau)$, write $\sigma^{\text{crys}}(\lambda, \tau) := L_{\underline{\lambda}} \otimes_{\mathcal{O}} \sigma^{\text{crys}, \circ}(\tau)$, and write $\bar{\sigma}^{\text{crys}}(\lambda, \tau)$ for the semisimplification of the \mathbf{F} -representation of $\text{GL}_d(k)$ given by $\sigma^{\text{crys}}(\lambda, \tau) \otimes_{\mathcal{O}} \mathbf{F}$. For each Serre weight \underline{k} , we write $F_{\underline{k}}$ for the corresponding irreducible \mathbf{F} -representation of $\text{GL}_d(k)$. Then there are unique integers $n_{\underline{k}}^{\text{crys}}(\lambda, \tau)$ such that

$$\bar{\sigma}^{\text{crys}}(\lambda, \tau) \cong \bigoplus_{\underline{k}} F_{\underline{k}}^{\oplus n_{\underline{k}}^{\text{crys}}(\lambda, \tau)}.$$

Our “universal” geometric Breuil–Mézard conjecture is as follows (see Lecture 9.4 below for some motivation for this conjecture).

9.2.5. Conjecture. *There are cycles $Z_{\underline{k}}$ with the property that for each regular Hodge type $\underline{\lambda}$ and each inertial type τ , we have $Z_{\text{crys}, \underline{\lambda}, \tau} = \sum_{\underline{k}} n_{\underline{k}}^{\text{crys}}(\lambda, \tau) \cdot Z_{\underline{k}}$.*

9.3. The relationship between the numerical, refined and geometric Breuil–Mézard conjectures. In brief (see [EG19, §8.3] for the details), the relationship is as follows: Conjecture 9.2.5 implies (by pulling back to versal rings) the geometric conjecture of [EG14], which in turn implies the numerical conjecture.

Conversely, the numerical conjecture implies Conjecture 9.2.5; in fact, it is enough to know the numerical conjecture for a single sufficiently generic $\bar{\rho}$ on each irreducible component of $\mathcal{X}_{d, \text{red}}$. To see this, recall that the numerical conjecture for $\bar{\rho}$ is that there are integers $\mu_{\underline{k}}(\bar{\rho})$ such that for all \underline{k} , we have

$$(9.3.1) \quad e(\text{Spec } R_{\bar{\rho}}^{\text{crys}, \lambda, \tau} / \varpi) = \sum_{\underline{k}} n_{\underline{k}}^{\text{crys}}(\lambda, \tau) \mu_{\underline{k}}(\bar{\rho}).$$

For each \underline{k} we choose a point $x_{\underline{k}} : \text{Spec } \mathbf{F} \rightarrow \overline{\mathcal{X}}_{d, \text{red}}$ which is contained in $\overline{\mathcal{X}}^{\underline{k}}$ and not in any $\overline{\mathcal{X}}^{\underline{k}'}$ for $\underline{k}' \neq \underline{k}$. We furthermore demand that $x_{\underline{k}}$ is a smooth point of $\overline{\mathcal{X}}_{d, \text{red}}$. (Since $\overline{\mathcal{X}}_{d, \text{red}}$ is reduced and of finite type over \mathbf{F} , there is a dense set of points of $\overline{\mathcal{X}}^{\underline{k}}$ satisfying these conditions.) Write $\bar{\rho}_{\underline{k}} : G_K \rightarrow \text{GL}_d(\mathbf{F})$ for

the representation corresponding to $X_{\underline{k}}$, and assume that the numerical conjecture holds for each $\bar{\rho}_{\underline{k}}$. Then if we set

$$(9.3.2) \quad Z_{\underline{k}} := \sum_{\underline{k}'} \mu_{\underline{k}}(\bar{\rho}_{\underline{k}'}) \cdot \bar{\chi}^{\underline{k}'},$$

it is easy to check that Conjecture 9.2.5 holds.

The numerical Breuil–Mézard conjecture (and consequently Conjecture 9.2.5) holds if $K = \mathbf{Q}_p$ and $d = 2$. Most cases are proved in Kisin’s paper [Kis09] and Paškūnas’ paper [Paš15], and the remaining cases not handled by these papers are proved in the papers [HT15, San14, Tun18, Tun19]. The conjecture also holds if $d = 2$, $p > 2$, $\lambda = (\underline{0}, \underline{1})$, and K and τ are arbitrary, by the main result of [GK14]. For some recent progress for $d > 2$ in suitably generic situations, we refer the reader to [LHLM20].

In the case that $d = 2$, $p > 2$, and K is arbitrary, we can make the cycles $Z_{\underline{k}}$ completely explicit: we say that a Serre weight \underline{k} for GL_2 is “Steinberg” if for each $\bar{\sigma}$ we have $k_{\bar{\sigma},1} - k_{\bar{\sigma},2} = p - 1$. If \underline{k} is Steinberg then we define $\tilde{\underline{k}}$ by $\tilde{k}_{\bar{\sigma},1} = \tilde{k}_{\bar{\sigma},2} = k_{\bar{\sigma},2}$. Then if \underline{k} is not Steinberg, we have $Z_{\underline{k}} = \mathcal{X}_2^{\underline{k}}$, while if \underline{k} is Steinberg, $Z_{\underline{k}} = \mathcal{X}_2^{\tilde{\underline{k}}} + \mathcal{X}_2^{\underline{k}}$.

This explicit description follows from the results of [CEGS19]; for the details, see [EG19, Thm. 8.6.2]. (Roughly speaking, the point is that it’s easy to compute the tamely potentially Barsotti–Tate deformation rings for generic extensions of generic characters, and they’re either zero or formally smooth.)

9.4. The weight part of Serre’s conjecture. Finally, we very briefly explain some motivation for Conjecture 9.2.5, and its relationship to the weight part of Serre’s conjecture. This connection was first explained in the context of the numerical Breuil–Mézard conjecture in [GK14]. For more details, see for example [GHS18] (particularly Sections 3 and 4). We will further expand on this discussion in the context of a hypothetical p -adic local Langlands correspondence involving sheaves on \mathcal{X}_d in Lecture 10.3.

We expect that the cycles $Z_{\underline{k}}$ will be effective, in the sense that they are combinations of the $\mathcal{X}_d^{\underline{k}}$ with non-negative coefficients. Indeed, we expect that after pulling back to the special fibre of the universal deformation ring at some $\bar{\rho}$, the cycle $Z_{\underline{k}}$ will be precisely the support of the Taylor–Wiles–Kisin patched modules of mod p automorphic forms of weight \underline{k} , and the support of a sheaf is an effective cycle by definition. (Again, see for example [GHS18, §3, 4] for more details; we are slightly simplifying the situation by supposing that the Taylor–Wiles–Kisin method does not introduce any patching variables, but these are essentially irrelevant for our discussion in any case.)

The *weight part of Serre’s conjecture* is a prediction that the possible Serre weights of mod p automorphic forms giving rise to a fixed global Galois representation only depends on the restrictions of this representation to decomposition groups of places above p . If this is the case, there is a way to associate a corresponding set of Serre weights to each $\bar{\rho} : G_K \rightarrow \mathrm{GL}_d(\mathbf{F}_p)$. Such a description was originally given by Serre in the case $K = \mathbf{Q}_p$ and $d = 2$; while Serre’s recipe is completely explicit, in general it seems to be unreasonable to hope for such a description.

However, if we admit the standard expectation (which is closely related to the Fontaine–Mazur conjecture) that all irreducible components of the generic fibres

of crystalline deformation rings are witnessed by automorphic forms, it follows formally from the Taylor–Wiles–Kisin method that the cycle $Z_{\underline{k}}$ is indeed (as conjectured above) the support of the corresponding patched module of weight \underline{k} modular forms, and from this one easily deduces that representation $\bar{\rho}$ admits \underline{k} as a Serre weight if and only if $Z_{\underline{k}}$ is supported at $\bar{\rho}$.

Equivalently, we can rephrase the weight part of Serre’s conjecture in the following way: to each irreducible component of $\mathcal{X}_{d,\text{red}}$, we assign the set of weights \underline{k} with the property that $Z_{\underline{k}}$ is supported on this component. Then for each $\bar{\rho}$, the corresponding set of Serre weights is simply the union of the sets of weights for the irreducible components of $\mathcal{X}_{d,\text{red}}$ which contain $\bar{\rho}$.

We expect that in particular the irreducible component $\mathcal{X}_{d,\text{red}}^{\underline{k}}$ is assigned the Serre weight \underline{k} , but that the list of Serre weights associated to this irreducible component can be longer; indeed, it follows from the description given above of the cycles $Z_{\underline{k}}$ in the case $d = 2$ that if $k_{\bar{\sigma},1} - k_{\bar{\sigma},2} = 0$ for all $\bar{\sigma}$, then there are two Serre weights associated to $\mathcal{X}_{d,\text{red}}^{\underline{k}}$, namely \underline{k} and the corresponding Steinberg weight.

LECTURE 10. MODULI SPACES AND BERNSTEIN CENTERS

In this final lecture we discuss some of the general theory of moduli spaces associated to algebraic stacks, and explain how it applies to the stacks we have constructed in the previous lectures. We will see that there is a surprising connection between these ideas and the p -adic local Langlands correspondence.

10.1. Moduli spaces. An algebraic stack \mathcal{X} has an underlying “Zariski” topological space, formed by taking a smooth cover by a scheme, taking its topological space, and then taking a quotient topological space.

10.1.1. *Example.* If $\mathcal{X} = [\mathbf{A}^1/\mathbf{G}_m]$ (over some field k , with $t \in \mathbf{G}_m$ acting on $x \in \mathbf{A}^1$ via multiplication, i.e. $t \cdot x = tx$), then the associated topological space consists of two points, one of which is open and specializes to the other, which is closed. These two points correspond to the two orbits of \mathbf{G}_m acting on \mathbf{A}^1 — the open orbit and the closed orbit. The closure relation between the points corresponds to the closure relation between these orbits.

The topological space underlying \mathcal{X} is *just* a topological space, though; it doesn’t come equipped with a structure sheaf making it into a scheme. So we can ask: is there a morphism $f : \mathcal{X} \rightarrow X$ to X a scheme, or, more generally, to an algebraic space, which is the “best possible approximation to the stack \mathcal{X} ”? Somewhat more precisely, one might call X a *moduli space associated to \mathcal{X}* if f is initial in the space of maps to algebraic spaces. Or one could say that f is an *associated moduli space morphism*.

Unfortunately, this doesn’t really help you access X or say anything concrete about it, or even determine if such an X exists. In [Alp14, Prop. 7.1.1], Jared Alper gives properties that ensure that f is initial for morphisms to *locally separated* algebraic spaces. The properties are:

- (1) If k is algebraically closed, the induced map $\mathcal{X}(k) \rightarrow X(k)$ identifies the target with the quotient of the source by the equivalence relation generated by $x \sim y$ if $\overline{\{x\}} \cap \overline{\{y\}} \neq \emptyset$.

- (2) f is a universal submersion.⁵
- (3) $f_*\mathcal{O}_{\mathcal{X}} = \mathcal{O}_X$.

As Alper notes, the first property says that X has the right points, the second that it has the right topology, and the third that it has the right functions.

10.1.2. *Example.* In the case of $\mathcal{X} = [\mathbf{A}^1/\mathbf{G}_m]$ over the algebraically closed field k , there are two k -points: the origin, and the open point (corresponding to the open orbit of \mathbf{G}_m on \mathbf{A}^1). Since the open point specializes to the closed point, the associated moduli space is just the single point $\text{Spec } k$.

This still leaves open the question of when such an X exists. The theorem of Keel–Mori [KM97] (plus various technical improvements) implies that if \mathcal{X} is Deligne–Mumford (see Definition 2.5.1) then there exists X which is a “coarse moduli space” (the map $\mathcal{X} \rightarrow X$ induces a bijection on k -points for algebraically closed fields k). Unfortunately, this result doesn’t apply in our context: firstly, our moduli stacks of étale (φ, Γ) -modules are formal algebraic stacks, rather than being actually algebraic; but, more significantly, they are not Deligne–Mumford: their points have infinite automorphism groups (every Galois representation $\bar{\rho}$ admits at least the scalar matrices as automorphisms).

In the context of more general (i.e. not-necessarily Deligne–Mumford) algebraic stacks, Alper [Alp13, Alp14] has developed a theory of “good” and “adequate” moduli spaces. His theory is related to earlier ideas in Geometric Invariant Theory (GIT) — the difference between “good” and “adequate” (the latter is the more general notion) is related to the difference of behaviour of the representation theory of reductive linear algebraic groups in positive characteristic versus in characteristic zero. One feature of this theory is that if $f : \mathcal{X} \rightarrow X$ is good or adequate, then it is universally closed. Also, if \mathcal{X} admits a good or adequate moduli space, then all stabilizers at closed k -points are reductive. Further, the relation $x \sim y$ if $\overline{\{x\}} \cap \overline{\{y\}} \neq \emptyset$ is in fact an equivalence relation.

10.1.3. *Example.* If G is reductive over k and A is a finite type k -algebra with a G -action, then $[\text{Spec } A/G] \rightarrow \text{Spec } A^G$ is adequate (or good in characteristic 0). People often refer to $\text{Spec } A^G$ as the GIT quotient of $\text{Spec } A$ by G .

10.1.4. *Example.* Take \mathbf{P}^1 , with \mathbf{G}_m acting via scaling. This is similar to the example of \mathbf{G}_m acting on \mathbf{A}^1 considered above, but now there are *two* closed orbits (each of the points 0 and ∞), and the open orbit specializes to both of them. Thus the relation $x \sim y$ if $\overline{\{x\}} \cap \overline{\{y\}} \neq \emptyset$ is *not* an equivalence relation in this example, and so $[\mathbf{P}^1/\mathbf{G}_m]$ does not admit an adequate moduli space.

In Galois deformation theory, the passage from Galois representations to pseudorepresentations is related to the ideas that we’re discussing. One subtlety in the theory of pseudorepresentations is that traces don’t know about extension classes, so the passage to pseudorepresentations factors through semisimplification; indeed, two representations $\rho_1, \rho_2 : G_K \rightarrow \text{GL}_d(\overline{\mathbf{F}}_p)$ have the same underlying pseudorepresentation precisely if $\rho_1^{\text{ss}} = \rho_2^{\text{ss}}$. The d -dimensional pseudorepresentations of G_K over $\overline{\mathbf{F}}_p$ are then in bijection with the semisimple $\rho : G_K \rightarrow \text{GL}_d(\overline{\mathbf{F}}_p)$.

⁵A *submersion* is a morphism which is surjective and which induces a quotient map on the underlying topological spaces.

These phenomena mirror corresponding phenomena related to the points of $\mathcal{X}_{d,\text{red}}$. Namely, the closed points of $\mathcal{X}_{d,\text{red}}(\overline{\mathbf{F}}_p)$ correspond to semisimple representations, and specialization to a closed point corresponds to semisimplification. That is, if ρ is any element of $\mathcal{X}_{d,\text{red}}(\overline{\mathbf{F}}_p)$, then the closure $\overline{\{\rho\}}$ has a unique closed point, which is ρ^{ss} .

Thus if we could find a moduli space map $\mathcal{X} \rightarrow X$ (whatever its precise meaning in the context of formal algebraic stacks), we might expect X to be a moduli space of pseudorepresentations. This is one of our main motivations for studying associated moduli spaces in our context.

The fact that any point of $\mathcal{X}_{d,\text{red}}(\overline{\mathbf{F}}_p)$ specializes to a unique closed point implies that the relation $x \sim y$ considered above is an equivalence relation. Since closed points of $\mathcal{X}_{d,\text{red}}$ correspond to semisimple representations, we also see that the stabilizers of closed points are reductive. These are two of the properties that are necessary to admit an adequate moduli space. Nevertheless, we will see that $\mathcal{X}_{d,\text{red}}$ does not admit an adequate moduli space when $d > 1$, and so we seem to be outside the scope of any generally developed theory.

10.1.5. *Example.* Consider $[(\mathbf{A}^2 \setminus \{0\})/\mathbf{G}_m]$, with the \mathbf{G}_m -action given via $t \cdot (x, y) = (x, ty)$. Then the open orbits in each of the vertical lines satisfying $y \neq 0$ specialize to the intersection of the line with the horizontal axis, away from the origin. On the other hand, the vertical line $y = 0$ is a single closed orbit. There's an obvious map $[(\mathbf{A}^2 \setminus \{0\})/\mathbf{G}_m] \rightarrow \mathbf{A}^1$, given by $(x, y) \mapsto x$, which is the associated moduli space map (in that it satisfies Alper's properties (1), (2), (3) above; in fact, it is even initial for morphisms to arbitrary algebraic spaces). This map is *not* adequate, for instance because it's not closed.

10.1.6. *Remark.* The preceding example illustrates that in general, the quotient of a *quasi-affine* scheme by a reductive group can be much nastier than the quotient of an affine scheme by a reductive group.

As we saw in Lecture 1, stacks similar to that in the previous example appear as moduli stacks of two-dimensional Fontaine–Laffaille modules, and so also appear as irreducible components in $\mathcal{X}_{2,\text{red}}$. Thus $\mathcal{X}_{2,\text{red}}$ (and, more generally, $\mathcal{X}_{d,\text{red}}$ for $d > 1$) will not admit an adequate moduli space.

Nevertheless, we anticipate the following result.

10.1.7. **Expected Theorem ([DEG]).** *Let $K = \mathbf{Q}_p$, and fix a character $\psi : G_{\mathbf{Q}_p} \rightarrow \overline{\mathbf{Z}}_p^\times$. Then $\mathcal{X}_2^{\det=\psi}$ admits an associated formal algebraic moduli space X (in an appropriately understood sense), with X_{red} being a certain chain of \mathbf{P}^1 s. The points of $X(\overline{\mathbf{F}}_p)$ correspond to 2-dimensional pseudorepresentations of $G_{\mathbf{Q}_p}$ over $\overline{\mathbf{F}}_p$ with determinant ψ (equivalently, to semi-simple $\rho : G_{\mathbf{Q}_p} \rightarrow \text{GL}_2(\overline{\mathbf{F}}_p)$ with $\det \rho = \psi$). The complete local ring of X at one of its closed points is naturally identified with the corresponding pseudodeformation ring.*

10.1.8. *Example.* The preceding expected theorem fits nicely with the deformation theory of 2-dimensional crystalline representations of $G_{\mathbf{Q}_p}$. The \mathbf{Z}_p -points of one of the crystalline stacks we have constructed correspond to Galois-invariant lattices in crystalline representations; but when we pass to pseudodeformations, the different possible lattices all share a common image.

In particular if we fix the Hodge–Tate weights to be $0, k - 1$ with $k \geq 2$, the images of these points in the moduli space X will be a family of crystalline representations V_{k, a_p} , parameterised by the crystalline Frobenius a_p with $|a_p| \leq 1$, i.e. parameterized by the rigid analytic closed unit disk. (That there is just a single parameter a_p is a simple calculation with weakly admissible modules.) The image of the crystalline moduli stack itself will then be a certain formal model of the closed unit disk, whose underlying reduced scheme embeds into the chain of \mathbf{P}^1 s given by X_{red} . Now the simplest formal model of the closed unit disk is $\widehat{\mathbf{A}}^1_{\mathbf{Z}_p}$, whose special fibre \mathbf{A}^1 certainly embeds into \mathbf{P}^1 . More general formal models are obtained by blowing up points in $\widehat{\mathbf{A}}^1_{\mathbf{Z}_p}$, and the special fibre then contains additional \mathbf{P}^1 s. In the case that $k \leq 2p + 1$, this can be seen explicitly by looking at the formulae for the reductions of crystalline representations in [Ber11, Thm. 5.2.1].

In fact, the results of [Ber11, Thm. 5.2.1] were an important clue that our stacks should exist. Indeed, Mark Kisin suggested in around 2004 that some kind of non-formal moduli (i.e. with $\bar{\rho}$ varying) spaces of local crystalline Galois representations should exist, motivated by the conjectures of [Bre03] and the results of [BB05].

10.2. Bernstein Centers. When we investigate the relationship between automorphic forms and Galois representations,⁶ the basic link between the two concepts is that the Hecke eigenvalues of an automorphic eigenform at primes not dividing the level should match with traces of Frobenius at primes that are unramified in the associated Galois representation.

At primes that divide the level, there is a more subtle way to extract eigenvalues from automorphic forms/automorphic representations. Namely, if K is (as always) a finite extension of \mathbf{Q}_p , then the abelian category of smooth representations of $\text{GL}_d(K)$ on \mathbf{C} -vector spaces admits a commutative ring of endomorphisms, called the *Bernstein centre* for $\text{GL}_d(K)$. On any irreducible representation, it will act via scalars (because of Schur’s Lemma). On unramified representations we recover the usual Hecke eigenvalues.

If f is an automorphic Hecke eigenform, for GL_d over some number field F , generating an automorphic representation π , and if K arises as the completion of F at some prime v above p , then π has a local factor π_v at v , which is a representation of $\text{GL}_d(K)$. This gives rise to a system of eigenvalues for the Bernstein centre. If v does not divide the level of f , this is just the usual collection of Hecke eigenvalues of f .

On the other hand, if $\rho : G_F \rightarrow \text{GL}_d(\overline{\mathbf{Q}}_\ell)$ is the global ℓ -adic Galois representation (provably in some cases, conjecturally in others) associated to f , with ℓ chosen so that $\ell \neq p$, then we can consider the restriction of ρ to a decomposition group at v , giving rise to a representation $\rho_v : G_K \rightarrow \text{GL}_d(\overline{\mathbf{Q}}_\ell)$. We know in some cases, and anticipate in general, that ρ_v and π_v are related via the local Langlands correspondence (once we fix an isomorphism $\iota : \mathbf{C} \xrightarrow{\sim} \overline{\mathbf{Q}}_\ell$).

On the other hand, the pseudorepresentation associated to ρ_v encodes the traces of ρ_v on all the elements of G_K , and so we have two collections of numbers associated to the local-at- v aspects of our situation: the eigenvalues of the Bernstein centre acting on π_v , and the pseudocharacter of ρ_v . It turns out (as a kind of numerical

⁶This is a huge topic, of course, and we introduce it here only to provide motivation, and in the most abbreviated way possible!

shadow of the local Langlands correspondence) that these numbers also determine one another.

In fact, there is even an integral version of this statement. Namely, we can consider smooth representations of $\mathrm{GL}_d(K)$ on \mathbf{Z}_ℓ -modules, and form the corresponding \mathbf{Z}_ℓ -Bernstein centre. And we can consider the moduli stacks $\mathcal{V}_Q = [\mathrm{Spec} A_Q / \mathrm{GL}_d]$ parameterizing local-at- v Weil–Deligne representations over \mathbf{Z}_ℓ -algebras, described in Lecture 1. This stack is the quotient of an affine scheme by a reductive group, and so the associated moduli space is just $\mathrm{Spec} A_Q^{\mathrm{GL}_d}$.

The following theorem of Helm–Moss encapsulates the manner in which numerical data extracted from Galois representations matches with eigenvalues of the Bernstein centre.

10.2.1. Theorem ([HM18]). $\varprojlim_Q A_Q^{\mathrm{GL}_d}$ is the \mathbf{Z}_ℓ -Bernstein center for $\mathrm{GL}_d(K)$.

10.2.2. Example. If g is an element of the Weil–Deligne group WD_K , then “trace of $\rho_v(g)$ ” gives an element of each $\mathrm{Spec} A_Q^{\mathrm{GL}_d}$, and thus an element h_σ of the Bernstein centre. The value of this function at a particular ρ_v — which is just the trace of the particular matrix $\rho_v(g)$ — will then correspond to the eigenvalue of h_σ on the representation π_v associated to ρ_v via local Langlands.

We can then ask if there is an $\ell = p$ analogue of this result. Even in the case of $\mathrm{GL}_2(\mathbf{Q}_p)$, Expected Theorem 10.1.7 shows the associated formal moduli space to the stack $(\mathcal{X}_2)^{\det=\psi}$ is *not* formally affine, so there is no obvious ring appearing on the Galois/ (φ, Γ) -module side to compare with a Bernstein centre. However, we have the following result.

10.2.3. Expected Theorem ([DEG]). If \mathcal{A} is the abelian category of smooth $\mathrm{GL}_2(\mathbf{Q}_p)$ -representations on \mathbf{Z}_p -modules which are locally p -power torsion, with central character equal to $\psi\varepsilon$, then \mathcal{A} localizes to a stack of categories over the formal scheme X of Expected Theorem 10.1.7.

Given this, we can form a *sheaf* of Bernstein centres of \mathcal{A} over X . Of course, we also have the structure sheaf \mathcal{O}_X . Our expectation then is that these two sheaves can be identified. Just as in the $\ell \neq p$ case, this identification will be mediated via the local Langlands correspondence — but now we will have to use the *p -adic local Langlands correspondence* for $\mathrm{GL}_2(\mathbf{Q}_p)$.

Indeed, if $D \rightarrow \mathcal{X}_2^{\det=\chi}$ is the universal (φ, Γ) -module, then $(D \boxtimes \mathbf{P}^1) / (D^{\natural} \boxtimes \mathbf{P}^1)$ (as defined by Colmez [Col10]) should be a quasi-coherent sheaf of $\mathrm{GL}_2(\mathbf{Q}_p)$ -representations over \mathcal{X}_2 . Then the identification between \mathcal{O}_X and the sheaf of Bernstein centres of \mathcal{A} will be determined by saying that the two sheaves of rings act in the same way on this family.

This is all strongly related to, and uses, Paškūnas’ work [Paš13].

10.3. A conjectural p -adic local Langlands correspondence. We return to the case of general K and d , and end our lectures with a vague and speculative discussion of a possible p -adic local Langlands correspondence, analogous to the conjectures made in the $\ell \neq p$ setting in [BZCHN20, Hel20, Zhu20]. We also explain briefly how such a conjecture could explain the geometric Breuil–Mézard conjecture of the previous lecture.

Each of the papers [BZCHN20, Hel20, Zhu20] proposes a certain conjectural enhancement of Theorem 10.2.1. We state a rough form of this conjecture here.

10.3.1. Rough Conjecture. *The category of $\mathrm{GL}_d(K)$ -representations on \mathbf{Z}_ℓ -modules admits a fully faithful embedding into the category of coherent sheaves on the Ind-algebraic stack $\varinjlim_Q \mathcal{V}_Q$, compatibly with the identification of Theorem 10.2.1.*

We have omitted all technical details and caveats from this statement (in particular, it should probably be stated at a derived level). The key intuition, though, is that we are conjecturing that we can localize $\mathrm{GL}_d(K)$ -representations in some fashion as sheaves of \mathcal{O} -modules over the moduli stack of rank d Weil–Deligne representations. And although the action of the Bernstein centre is not sufficient to separate all irreducible representations (e.g. any two such representations that admit a non-trivial extension, such as the trivial and Steinberg representations of $\mathrm{GL}_2(K)$, will have the same system of eigenvalues under the Bernstein centre), this more refined localization process should be able to distinguish non-isomorphic irreducibles.

There is another important facet of the (conjectural) fully faithful embedding of Conjecture 10.3.1 which we want to explain, if only in somewhat general and vague terms. The essential point is that there is *another* method of obtaining coherent sheaves out of $\mathrm{GL}_d(K)$ -representations, not on the stacks \mathcal{V}_Q themselves, but on their versal rings at closed points, which is to say, on formal deformation rings of Galois representations. This is the method of *Taylor–Wiles–Kisin* patching (which already made an appearance in the previous lecture).

Briefly, if we choose some global context related to $\mathrm{GL}_d(K)$ (e.g. a modular curve or Shimura curve for GL_2 , or a certain unitary Shimura variety more generally), and an irreducible⁷ automorphic global Galois representation $r : G_F \rightarrow \mathrm{GL}_d(\overline{\mathbf{F}}_p)$ (here F is a totally real or CM field of which K is a certain completion, and which is related to the Shimura variety at hand), and an ℓ -adic representation σ of $\mathrm{GL}_d(\mathcal{O}_K)$, then we are able to “patch” appropriate Hecke localizations of the cohomology of whichever Shimura variety is in play so as to obtain a coherent sheaf, traditionally denoted by $M_\infty(\sigma)$, which lives over the formal neighbourhood⁸ of $r|_{G_K}$.

The relationship between patching and Conjecture 10.3.1 can be stated very roughly as follows:⁹ In sufficiently good situations,¹⁰ $M_\infty(\sigma)$ should be equal to the formal completion at $r|_{G_K}$ of the coherent sheaf associated to $c\text{-Ind}_{\mathrm{GL}_d(\mathcal{O}_K)}^{\mathrm{GL}_d(K)} \sigma$ by Conjecture 10.3.1. In general, the patched module $M_\infty(\sigma)$ will be related to this latter coherent sheaf, but we will have to take into account level structure at completions of F other than K , the most subtle being the level structure at the primes above ℓ ; and this latter case amounts to studying patching for p -adic $\mathrm{GL}_n(\mathcal{O}_K)$ -representations.

The patching described above, in the $\ell \neq p$ case (see e.g. [Sho18, Man20, MS20]), is less commonly studied than the case of $\ell = p$, and we now return to the latter

⁷In fact r should satisfy a stronger condition than irreducibility, but we will suppress all such technical details.

⁸We are here ignoring the “patching variables” that might have been introduced to effect the patching process; these *should* be ignorable, although this is not known in general.

⁹For more precise statements, the reader can look at [Zhu20, Conj. 4.6.3], which treats the function field analogue of the situation we are describing here — although the connection to patching is not made precise there. The case of number fields should be discussed carefully (just at the level of conjecture, to be sure!) in forthcoming work of Xinwen Zhu and ME.

¹⁰E.g. if p is unramified in F , and if we choose the level and coefficients of our cohomology carefully enough.

case. In this context, the appropriate stacks over which to state an analogue of Conjecture 10.3.1 are our moduli stacks \mathcal{X}_d parameterizing étale (φ, Γ) -modules, in place of the moduli stacks of Weil–Deligne representations.

Our hope, then, is that there is a canonically defined quasi-coherent sheaf Π of $\mathrm{GL}_d(K)$ -representations on \mathcal{X}_d which in particular enjoys the properties that it satisfies local-global compatibility for completed cohomology of locally symmetric spaces (in the sense of being compatible with Taylor–Wiles–Kisin patching, as carried out in e.g. [CEG⁺16]), and explains the Breuil–Mézard conjecture and the weight part of Serre’s conjecture in a sense that we will explain shortly.

The analogy with Conjecture 10.3.1 comes by considering the functor $\pi \mapsto \mathrm{RHom}_{\mathrm{GL}_d(K)}(\pi, \Pi)$ from the derived category of smooth representations of $\mathrm{GL}_d(K)$ on \mathbf{Z}_p -algebras to the derived category of quasi-coherent sheaves on \mathcal{X}_d ; this functor should be fully faithful. This being the case, one could in particular localize the former derived category over \mathcal{X}_d , generalizing the case $d = 2$, $K = \mathbf{Q}_p$ explained above. (However, in this more general context, there does not seem to be any obvious analogue of the moduli space for \mathcal{X}_d that we considered above, and it seems likely that one really has to consider this localization as taking place over the stack.)

In the case that $d = 2$ and $K = \mathbf{Q}_p$, we explicitly construct a candidate quasi-coherent sheaf Π in [DEG], by generalising Colmez’s construction from [Col10] of the representation he denotes $(D \boxtimes \mathbf{P}^1)/(D^{\natural} \boxtimes \mathbf{P}^1)$. Note that after pulling back to versal rings, we know that this sheaf realises the p -adic local Langlands correspondence for $\mathrm{GL}_2(\mathbf{Q}_p)$ by results of Colmez and Paškūnas [Paš13] (see also [CEG⁺18] for a perspective related to patching). In particular, a construction of Π in general would generalize the p -adic local Langlands correspondence from the case of $\mathrm{GL}_2(\mathbf{Q}_p)$ to the general case of $\mathrm{GL}_d(K)$.

If σ is a representation of $\mathrm{GL}_n(\mathcal{O}_K)$ on a finitely generated \mathbf{Z}_p -module, then we can apply the functor $\mathrm{RHom}_{\mathrm{GL}_d(K)}(-, \Pi)$ to $\pi := c\text{-Ind}_{\mathrm{GL}_d(\mathcal{O}_K)}^{\mathrm{GL}_d(K)} \sigma^{\vee}$. (The Pontrjagin dual is included just so as to make the functor covariant.) We anticipate that this RHom should be supported in degree 0, and that the resulting sheaf $\mathcal{M}(\sigma) := \mathrm{Hom}_{\mathrm{GL}_d(\mathbf{Z}_p)}(\sigma^{\vee}, \Pi)$ will be coherent. (Just as in the $\ell \neq p$ case discussed above, the formal completions of $\mathcal{M}(\sigma)$ at $\overline{\mathbf{F}}_p$ -points of $\mathcal{X}_{d, \mathrm{red}}$ should give the usual patched modules $M_{\infty}(\sigma)$.) These (conjectural) coherent sheaves allow us to connect the present discussion to the Breuil–Mézard and Serre weight conjectures.

Indeed, we expect that if k is a Serre weight (regarded as usual as a representation of $\mathrm{GL}_d(\mathcal{O}_K)$ via inflation from $\mathrm{GL}_d(k)$), then the cycle Z_k considered in Lecture 9 is simply the support of $\mathcal{M}(F_k)$. Furthermore, we expect that if λ is a regular Hodge type, τ is an inertial type, and $\sigma^{\circ}(\lambda, \tau)$ is the representation considered in Lecture 9, then the support of $\mathcal{M}(\sigma^{\circ}(\lambda, \tau))$ is exactly $\mathcal{X}_d^{\mathrm{crys}, \lambda, \tau}$. This being the case, Conjecture 9.2.5 is an immediate consequence of the additivity of supports, while the expectation that the cycles Z_k encode the weight part of Serre’s conjecture should be a consequence of a local-global compatibility for completed cohomology of Shimura varieties (or more generally locally symmetric spaces), generalising the results of [Eme11, CEG⁺18] for the modular curve.

This may all seem a speculation too far, but we note that while there is in general no candidate for Π , the construction of [CEG⁺16] does produce such a candidate after pulling back to the versal deformation rings; and the conjectural explanation above for the Breuil–Mézard conjecture describes exactly how the results of [Kis09,

[GK14] are proved (in the setting of the numerical conjecture, over a deformation ring).

REFERENCES

- [All19] Patrick B. Allen, *On automorphic points in polarized deformation rings*, American Journal of Mathematics **141** (2019), no. 1, 119–167.
- [Alp13] Jarod Alper, *Good moduli spaces for Artin stacks*, Ann. Inst. Fourier (Grenoble) **63** (2013), no. 6, 2349–2402.
- [Alp14] ———, *Adequate moduli spaces and geometrically reductive group schemes*, Algebr. Geom. **1** (2014), no. 4, 489–531.
- [Ax70] James Ax, *Zeros of polynomials over local fields—The Galois action*, J. Algebra **15** (1970), 417–428.
- [BB05] Laurent Berger and Christophe Breuil, *Sur la réduction des représentations cristallines de dimension 2 en poids moyens*, Available at <https://arxiv.org/abs/math/0504388>.
- [Ber11] Laurent Berger, *La correspondance de Langlands locale p -adique pour $GL_2(\mathbf{Q}_p)$* , Astérisque (2011), no. 339, Exp. No. 1017, viii, 157–180, Séminaire Bourbaki. Vol. 2009/2010. Exposés 1012–1026.
- [BLGGT14] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor, *Potential automorphy and change of weight*, Ann. of Math. (2) **179** (2014), no. 2, 501–609.
- [BMS18] Bhargav Bhatt, Matthew Morrow, and Peter Scholze, *Integral p -adic Hodge theory*, Publ. Math. Inst. Hautes Études Sci. **128** (2018), 219–397.
- [Bre03] Christophe Breuil, *Sur quelques représentations modulaires et p -adiques de $GL_2(\mathbf{Q}_p)$. II*, J. Inst. Math. Jussieu **2** (2003), no. 1, 23–58.
- [BZCHN20] David Ben-Zvi, Harrison Chen, David Helm, and David Nadler, *Coherent Springer theory and the categorical Deligne–Langlands correspondence*, 2020.
- [CEG⁺16] Ana Caraiani, Matthew Emerton, Toby Gee, David Geraghty, Vytautas Paškūnas, and Sug Woo Shin, *Patching and the p -adic local Langlands correspondence*, Camb. J. Math. **4** (2016), no. 2, 197–287.
- [CEG⁺18] ———, *Patching and the p -adic Langlands program for $GL_2(\mathbf{Q}_p)$* , Compos. Math. **154** (2018), no. 3, 503–548.
- [CEGS19] Ana Caraiani, Matthew Emerton, Toby Gee, and David Savitt, *Moduli stacks of two-dimensional Galois representations*, arXiv e-prints (2019), arXiv:1908.07019.
- [CL11] Xavier Caruso and Tong Liu, *Some bounds for ramification of p^n -torsion semi-stable representations*, J. Algebra **325** (2011), 70–96.
- [Col10] Pierre Colmez, *Représentations de $GL_2(\mathbf{Q}_p)$ et (ϕ, Γ) -modules*, Astérisque (2010), no. 330, 281–509.
- [Dee01] Jonathan Dee, *Φ - Γ modules for families of Galois representations*, J. Algebra **235** (2001), no. 2, 636–664.
- [DEG] Andrea Dotto, Matthew Emerton, and Toby Gee, *In preparation*.
- [DHKM20] Jean-François Dat, David Helm, Robert Kurinczuk, and Gilbert Moss, *Moduli of Langlands Parameters*, 2020.
- [Dri06] Vladimir Drinfeld, *Infinite-dimensional vector bundles in algebraic geometry: an introduction*, The unity of mathematics, Progr. Math., vol. 244, Birkhäuser Boston, Boston, MA, 2006, pp. 263–304.
- [EG14] Matthew Emerton and Toby Gee, *A geometric perspective on the Breuil–Mézard conjecture*, J. Inst. Math. Jussieu **13** (2014), no. 1, 183–223.
- [EG19] Matthew Emerton and Toby Gee, *Moduli stacks of étale (φ, Γ) -modules and the existence of crystalline lifts*, arXiv e-prints (2019), arXiv:1908.07185.
- [EG21] Matthew Emerton and Toby Gee, *“Scheme-theoretic images” of certain morphisms of stacks*, Algebraic Geometry **8** (2021), no. 1, 1–132.
- [Eme] Matthew Emerton, *Formal algebraic stacks*, Available at <http://www.math.uchicago.edu/~emerton/pdffiles/formal-stacks.pdf>.
- [Eme11] ———, *Local-global compatibility in the p -adic Langlands programme for GL_2/\mathbf{Q}* , 2011.
- [FGK11] Kazuhiro Fujiwara, Ofer Gabber, and Fumiharu Kato, *On Hausdorff completions of commutative rings in rigid geometry*, J. Algebra **332** (2011), 293–321.

- [FK18] Kazuhiro Fujiwara and Fumiharu Kato, *Foundations of rigid geometry. I*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2018.
- [Gao19] Hui Gao, *Breuil–Kisin modules and integral p -adic Hodge theory*, arXiv e-prints (2019), arXiv:1905.08555.
- [GHS18] Toby Gee, Florian Herzig, and David Savitt, *General Serre weight conjectures*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 12, 2859–2949.
- [GK14] Toby Gee and Mark Kisin, *The Breuil–Mézard conjecture for potentially Barsotti–Tate representations*, Forum Math. Pi **2** (2014), e1 (56 pages).
- [Hel20] Eugen Hellmann, *On the derived category of the Iwahori–Hecke algebra*, 2020.
- [Her98] Laurent Herr, *Sur la cohomologie galoisienne des corps p -adiques*, Bull. Soc. Math. France **126** (1998), no. 4, 563–600.
- [HM18] David Helm and Gilbert Moss, *Converse theorems and the local Langlands correspondence in families*, Invent. Math. **214** (2018), no. 2, 999–1022.
- [HT15] Yongquan Hu and Fucheng Tan, *The Breuil–Mézard conjecture for non-scalar split residual representations*, Ann. Sci. Éc. Norm. Supér. (4) **48** (2015), no. 6, 1383–1421.
- [Kis06] Mark Kisin, *Crystalline representations and F -crystals*, Algebraic geometry and number theory, Progr. Math., vol. 253, Birkhäuser Boston, Boston, MA, 2006, pp. 459–496.
- [Kis08] ———, *Potentially semi-stable deformation rings*, J. Amer. Math. Soc. **21** (2008), no. 2, 513–546.
- [Kis09] ———, *The Fontaine–Mazur conjecture for GL_2* , J. Amer. Math. Soc. **22** (2009), no. 3, 641–690.
- [KM97] Seán Keel and Shigefumi Mori, *Quotients by groupoids*, Ann. of Math. (2) **145** (1997), no. 1, 193–213.
- [KPX14] Kiran S. Kedlaya, Jonathan Pottharst, and Liang Xiao, *Cohomology of arithmetic families of (φ, Γ) -modules*, J. Amer. Math. Soc. **27** (2014), no. 4, 1043–1115.
- [LHLM20] Daniel Le, Bao V. Le Hung, Brandon Levin, and Stefano Morra, *Local models for Galois deformation rings and applications*, 2020.
- [Liu18] Tong Liu, *Compatibility of Kisin modules for different uniformizers*, J. Reine Angew. Math. **740** (2018), 1–24.
- [Man20] Jeffrey Manning, *Patching and Multiplicity 2^k for Shimura Curves*, 2020.
- [MS20] Jeff Manning and Jack Shotton, *Ihara’s lemma for Shimura curves over totally real fields via patching*, 2020.
- [Nek93] Jan Nekovář, *On p -adic height pairings*, Séminaire de Théorie des Nombres, Paris, 1990–91, Progr. Math., vol. 108, Birkhäuser Boston, Boston, MA, 1993, pp. 127–202.
- [Paš13] Vytautas Paškūnas, *The image of Colmez’s Montreal functor*, Publ. Math. Inst. Hautes Études Sci. **118** (2013), 1–191.
- [Paš15] ———, *On the Breuil–Mézard conjecture*, Duke Math. J. **164** (2015), no. 2, 297–359.
- [PR09] G. Pappas and M. Rapoport, *Φ -modules and coefficient spaces*, Mosc. Math. J. **9** (2009), no. 3, 625–663, back matter.
- [San14] Fabian Sander, *Hilbert–Samuel multiplicities of certain deformation rings*, Math. Res. Lett. **21** (2014), no. 3, 605–615.
- [Sho18] Jack Shotton, *The Breuil–Mézard conjecture when $l \neq p$* , Duke Math. J. **167** (2018), no. 4, 603–678.
- [Sta] The Stacks Project Authors, *Stacks Project*, <http://stacks.math.columbia.edu>.
- [SZ99] P. Schneider and E.-W. Zink, *K -types for the tempered components of a p -adic general linear group*, J. Reine Angew. Math. **517** (1999), 161–208, With an appendix by Schneider and U. Stuhler.
- [Tun18] S.-N. Tung, *On the automorphy of 2-dimensional potentially semi-stable deformation rings of $G_{\mathbf{Q}_p}$* , ArXiv e-prints (2018).
- [Tun19] Shen-Ning Tung, *On the modularity of 2-adic potentially semi-stable deformation rings*, arXiv e-prints (2019), arXiv:1908.06174.
- [Zhu20] Xinwen Zhu, *Coherent sheaves on the stack of Langlands parameters*, 2020.

Email address: emerton@math.uchicago.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVE.,
CHICAGO, IL 60637

Email address: `toby.gee@imperial.ac.uk`

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON, LONDON SW7 2AZ, UK