Chapter III

Galois Representations

The focus of this chapter is the construction of Galois representations associated to p-adic modular forms. Specifically, we shall be interested the problem of studying the "deformations of a residual Galois representation" as considered by Mazur in [Ma]. Thus, we will begin with an absolutely irreducible Galois representation defined over a finite field, which we will assume to be attached, as in the work of Deligne and Serre, to a modular form defined over that field. We will then consider its liftings to complete noetherian overrings. In [Ma], Mazur constructed a universal lifting of this kind, and studied its properties in some detail. We will show that, under the assumption that the residual representation is attached to a modular form, a good portion of the liftings he obtains are in fact attached to p-adic modular forms.

The fact that one can attach p-adic representations to p-adic modular forms was first noticed by Hida, who, in [Hi86b] and [Hi86a], constructed analytic families of such representations attached to analytic families of ordinary p-adic modular forms; he also showed how to obtain a large number of such analytic families. This was further studied, still in the ordinary case, by Mazur and Wiles in [MW86], who constructed what may be called "the universal ordinary modular deformation", i.e., a family of deformations of a representation (which is assumed absolutely irreducible and attached to an ordinary modular form) which parametrizes all possible deformations attached to ordinary p-adic modular forms.

In this chapter, we continue in the spirit of these results, by constructing the universal modular deformation of the given (modular, absolutely irreducible) residual Galois representation, parametrizing all the deformations that are attached to p-adic modular forms of the given level. As in the case of Hida's work, the crucial step is to obtain a good theory of the duality between spaces of modular forms and their Hecke algebras, and we devote the first section of this chapter to constructing such a theory. We then outline a recent result of Hida which we feel should be better known, and proceed to consider the problem of constructing modular deformations of a given residual representation. Finally, we obtain an estimate on the dimension of the space of modular deformations, and consider its relation to the full space of deformations of the given residual representation. We conclude by formulating several questions that arise naturally from the theory as it is known at this point.

III.1 Duality Theorems

Our first goal is to study the duality between spaces of modular forms and their Hecke algebras. Our goal is to identify $\mathbf{V}$ as a certain space of functions on the Hecke algebra $\mathbf{T}$, and similarly, $\mathbf{V}_{par}$ as a certain space of functions on the corresponding Hecke algebra $\mathbf{T}_0$. This works perfectly well in the case of parabolic modular functions, but is a little more complicated in the general case because of the presence of the constants in $\mathbf{V}$. The result is obtained by starting from V. Miller's results in the finite-dimensional case, and using the inverse limit technique of the section on Hecke operators. These results are similar to results obtained by Hida in [Hi86b], except that Hida's results give a Pontryagin duality while ours give a duality of topological $\mathbf{Z}_p$-modules. In particular, we show that there is a bijection between (normalized) eigenforms in $\mathbf{V}_{par}$ and continuous $\mathbf{Z}_p$-algebra homomorphisms from $\mathbf{T}_0$ to p-adic rings.

III.1.1 Classical duality

To fix notation, let $\mathcal{H}_k(B,\mathrm{N}p^\nu)$ denote the Hecke algebra corresponding to the space $\mathsf{D}_k(B,\mathrm{N}p^\nu)$ of divided congruences of classical modular forms of weight at most k. We can define a bilinear form

$$\begin{array}{ccc} \mathcal{H}_k(B,\mathrm{N}p^\nu)\times \mathsf{D}_k(B,\mathrm{N}p^\nu) & \longrightarrow & B \\ (\mathrm{T},f) & \longmapsto & a_1(\mathrm{T}f), \end{array} \tag{III.1}$$

where $a_1(\mathrm{T}f)$ denotes the coefficient of q in the q-expansion of $\mathrm{T}f$. If we let $\mathsf{S}^k(B,\mathrm{N}p^\nu)\subset \mathsf{D}_k(B,\mathrm{N}p^\nu)$ denote the subspace of divided congruences of cusp forms , and let $\mathrm{h}_k(B,\mathrm{N}p^\nu)$ denote the corresponding Hecke algebra, then (III.1) clearly induces a bilinear form

$$\mathrm{h}_k(B,\mathrm{N}p^\nu)\times \mathsf{S}^k(B,\mathrm{N}p^\nu)\longrightarrow B.$$

Similarly, we get a bilinear form

$$\mathcal{H}'_k(B,\mathrm{N}p^\nu)\times \mathsf{D}'_k(B,\mathrm{N}p^\nu)\longrightarrow B,$$

where the primes have the same meaning as above, i.e., if K is the fraction field of B, then

$$\mathsf{D}'_k(K,\mathrm{N}p^\nu)=\bigoplus_{i=1}^{k} M(K,i,\mathrm{N}p^\nu),$$

$$\mathsf{D}'_k(B,\mathrm{N}p^\nu)=\{f\in \mathsf{D}'_k(K,\mathrm{N}p^\nu)|f(q)\in \mathbf{Z}_p[[q]],\}$$

and $\mathcal{H}'_k$ is the corresponding Hecke algebra. Then we have the following result:

Theorem III.1.1 *Let K be a finite extension of $\mathbf{Q}_p$, and let $\mathcal{O}_K$ be its ring of integers. Assume B is either K or $\mathcal{O}_K$, and let*

$$m_k(\mathcal{O}_K, \mathrm{N}p^\nu) = (K + \mathsf{D}_k(\mathcal{O}_K, \mathrm{N}p^\nu))/K.$$

Then the pairing (III.1) induces perfect pairings of B-modules

$$\begin{aligned} \mathcal{H}'_k(K, \mathrm{N}p^\nu) \times \mathsf{D}'_k(K, \mathrm{N}p^\nu) &\longrightarrow K \\ \mathcal{H}'_k(\mathcal{O}_K, \mathrm{N}p^\nu) \times m_k(\mathcal{O}_K, \mathrm{N}p^\nu) &\longrightarrow \mathcal{O}_K \\ \mathrm{h}_k(B, \mathrm{N}p^\nu) \times \mathsf{S}^k(B, \mathrm{N}p^\nu) &\longrightarrow B. \end{aligned}$$

Proof: This is [Hi86b, Proposition 2.1], where the result is attributed to V. Miller. □

One should note that Hida shows also that $\mathcal{H}'_k(\mathcal{O}_K, \mathrm{N}p^\nu)$ (which is defined as the Hecke algebra corresponding to $\mathsf{D}'_k(\mathcal{O}_K, \mathrm{N}p^\nu)$) "is" the Hecke algebra corresponding to $m_k(\mathcal{O}_K, \mathrm{N}p^\nu)$ (the subalgebra of the ring of endomorphisms generated by the Hecke and diamond operators); this follows from the (obvious) fact that

$$m_k(\mathcal{O}_K, \mathrm{N}p^\nu) \otimes K = \mathsf{D}'_k(\mathcal{O}_K, \mathrm{N}p^\nu) \otimes K = \mathsf{D}'_k(K, \mathrm{N}p^\nu) = \bigoplus_{i=1}^{k} M(K, i, \mathrm{N}p^\nu).$$

In particular, we have, setting $\mathsf{S}^k = \mathsf{S}^k(\mathbf{Z}_p, \mathrm{N}p^\nu)$ and $\mathrm{h}_k = \mathrm{h}_k(\mathbf{Z}_p, \mathrm{N}p^\nu)$, we get

$$\mathsf{S}^k \cong \mathrm{Hom}_{\mathbf{Z}_p}(\mathrm{h}_k, \mathbf{Z}_p)$$

(homomorphisms of $\mathbf{Z}_p$-modules), and it is easy to see that a cusp form $f \in \mathsf{S}^k$ corresponds to a homomorphism *of $\mathbf{Z}_p$-algebras* if and only if it is a normalized eigenform, that is, if and only if it is a simultaneous eigenform for all the Hecke and diamond operators and has $a_1(f) = 1$. Conversely, every algebra homomorphism $\mathrm{h}_k \longrightarrow \mathbf{Z}_p$ corresponds to a normalized eigenform (so that given a system of eigenvalues in $\mathbf{Z}_p$, there is a unique normalized eigenform belonging to it).

We would like to extend this to other p-adic rings B. For this, note that, since h_k is $\mathbf{Z}_p$-free,

$$\mathsf{S}^k \otimes B \cong \mathrm{Hom}_{\mathbf{Z}_p}(\mathrm{h}_k, \mathbf{Z}_p) \otimes B \cong \mathrm{Hom}_{\mathbf{Z}_p}(\mathrm{h}_k, B).$$

We will use this fact later, when we pass to the limit situation.

III.1.2 Duality for parabolic p-adic modular functions

Given the pairing on the finite-dimensional case, one may attempt to pass to the limit in order to obtain results for generalized p-adic modular functions. We consider first the case of parabolic modular functions, which is clearly simpler (because, as we saw above, the constants are a problem: for any constant $b \in B$ and any $\mathrm{T} \in \mathbf{T}$, we have $(\mathrm{T}, b) \mapsto 0$ under the pairing III.1)

Assume first that $B = \mathcal{O}_K$ as above, so that we have a perfect pairing

$$\mathfrak{h}_k(B, \mathrm{N}p^\nu) \times \mathsf{S}^k(B, \mathrm{N}p^\nu) \longrightarrow B.$$

The $\mathfrak{h}_k$ form an inverse system of $\mathbf{Z}_p$-algebras, while the S^k form a direct system of $\mathbf{Z}_p$-modules. Going to the limit, we get a pairing

$$\varprojlim_k \mathfrak{h}_k(B, \mathrm{N}p^\nu) \times \varinjlim_k \mathsf{S}^k(B, \mathrm{N}p^\nu) = \mathbf{T}_0(B, \mathrm{N}) \times \mathsf{S}(B, \mathrm{N}p^\nu) \longrightarrow B,$$

which gives maps

$$\mathbf{T}_0(B, \mathrm{N}) \xrightarrow{\sim} \mathrm{Hom}_B(\mathsf{S}(B, \mathrm{N}p^\nu), B) \xrightarrow{\sim} \mathrm{Hom}_B(\mathbf{V}_{par}(B, \mathrm{N}p^\nu), B),$$

and

$$\mathsf{S}(B, \mathrm{N}p^\nu) \longrightarrow \mathrm{Hom}_B(\mathbf{T}_0(B, \mathrm{N}), B).$$

It is clear that the first map, being the inverse limit of the maps

$$\mathfrak{h}_k(B, \mathrm{N}p^\nu) \xrightarrow{\sim} \mathrm{Hom}_B(\mathsf{S}^k(B, \mathrm{N}p^\nu), B),$$

is, as indicated, an isomorphism; the second isomorphism is immediate since S is p-adically dense in $\mathbf{V}_{par}$. As to the map

$$\mathsf{S} \longrightarrow \mathrm{Hom}_B(\mathbf{T}_0, B),$$

we know only that it is injective, since it is the direct limit of the maps

$$\mathsf{S}^k(B, \mathrm{N}p^\nu) \xrightarrow{\sim} \mathrm{Hom}_B(\mathfrak{h}_k(B, \mathrm{N}p^\nu), B).$$

It therefore identifies S with a submodule of $\mathrm{Hom}_B(\mathbf{T}_0(B, \mathrm{N}p^\nu), B)$, which is easily seen to be the submodule of all the B-module homomorphisms $f : \mathbf{T}_0 \longrightarrow B$ which factor through the projection $\mathbf{T}_0 \longrightarrow \mathfrak{h}_k$ for some k (since S is just the union of the S^k); any such homomorphism will be continuous if we give $\mathbf{T}_0$ its inverse limit topology (which makes it compact) and B its p-adic topology. Let us denote this submodule by $\mathrm{Hom}_B^{fact}(\mathbf{T}_0, B)$; it clearly depends on the particular representation of $\mathbf{T}_0$ as an inverse limit (as does S). The p-adic topology on S corresponds to the topology of uniform convergence on $\mathrm{Hom}_B^{fact}(\mathbf{T}_0, B)$ (i.e., to the sup norm induced by the p-adic norm on B). Since $\mathbf{V}_{par}$ is the p-adic completion of S, taking completions induces an identification between $\mathbf{V}_{par}$ and the completion of $\mathrm{Hom}_B^{fact}(\mathbf{T}_0, B)$; this last is contained in the submodule $\mathrm{Hom}_{B,conts}(\mathbf{T}_0, B)$ of *continuous* B-module homomorphisms (where $\mathbf{T}_0$ is given the inverse limit topology and B the p-adic topology), which is complete (with the topology of uniform convergence). Thus, we have obtained an inclusion $\mathbf{V}_{par} \hookrightarrow \mathrm{Hom}_{B,conts}(\mathbf{T}_0, B)$, mapping a parabolic modular function f to the homomorphism ϕ_f defined by $\phi_f(\mathrm{T}) = a_1(\mathrm{T}f)$.

Proposition III.1.2 *Let $B = \mathcal{O}_K$ be the ring of integers in a finite extension K of $\mathbf{Q}_p$, and let $\mathbf{T}_0(B, \mathrm{N})$ have the inverse limit topology, B and $\mathbf{V}_{par}(B, \mathrm{N})$ the p-adic topology, and $\mathrm{Hom}_{B,conts}(\mathbf{T}_0(B, \mathrm{N}), B)$ the topology of uniform convergence. Then the mapping*

$$\mathbf{V}_{par}(B, \mathrm{N}) \longrightarrow \mathrm{Hom}_{B,conts}(\mathbf{T}_0(B, \mathrm{N}), B)$$

defined by $f \mapsto \phi_f$ is an isomorphism of topological B-modules.

Proof: It suffices, after the above discussion, to show that any continuous homomorphism $\phi : \mathbf{T}_0 \longrightarrow B$ can be approximated by homomorphisms which factor through one of the h_k. Given ϕ, consider its reduction mod p^n,

$$\phi_n : \mathbf{T}_0 \longrightarrow B/p^nB.$$

Since B/p^nB is (finite and) discrete, it is clear that ϕ_n factors through some h_k, giving a map $\overline{\psi}_n : \mathrm{h}_k \to B/p^nB$, which can then be lifted to a map $\psi_n : \mathrm{h}_k \longrightarrow B$, because h_k is a free B-module. Then it is clear that $\psi_n \to \phi$, and we are done. □

The restriction to the case when $B = \mathcal{O}_K$ in the above result can be removed without too much trouble by using the fact that, for *any* p-adic ring B, $\mathbf{V}_{par}(B, \mathrm{N}) = \mathbf{V}_{par}(\mathbf{Z}_p, \mathrm{N})\hat{\otimes}B$, which reduces everything to the case of $\mathbf{Z}_p$. More generally, we can restrict to algebras over the Witt ring $W(\mathsf{k})$ of a finite field k, and get a general duality statement. Recall that we say an parabolic eigenform for the Hecke algebra $\mathbf{T}_0$ is *normalized* if the coefficient a_1 of q in its q-expansion is equal to 1. Then we have:

Corollary III.1.3 *Let k be a finite field, and let $W(\mathsf{k})$ be its ring of Witt vectors. For any p-adically complete $W(\mathsf{k})$-algebra B (with the p-adic topology), we have*

$$\mathbf{V}_{par}(B, \mathrm{N}) \cong \mathrm{Hom}_{W(\mathsf{k}),conts}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), B)$$

*(continuous homomorphisms of $W(\mathsf{k})$-modules), via the map $f \mapsto \phi_f$. Moreover, ϕ_f is a homomorphism of $W(\mathsf{k})$-*algebras *if and only if $f \in \mathbf{V}_{par}(B, \mathrm{N})$ is a normalized simultaneous eigenform for the Hecke and diamond operators. In particular, given any eigenform for $\mathbf{T}_0$, there exists a* normalized *eigenform with the same system of eigenvalues.*

Proof: For the first assertion,

$$\begin{aligned}
\mathbf{V}_{par}(B, \mathrm{N}) &= \varprojlim_n \mathbf{V}_{par}(W(\mathsf{k}), \mathrm{N}) \otimes B/p^nB \\
&= \varprojlim_n \mathsf{S}(W(\mathsf{k}), \mathrm{N}) \otimes B/p^nB \\
&\cong \varprojlim_n \mathrm{Hom}_{W(\mathsf{k}),conts}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), \mathbf{Z}_p) \otimes B/p^nB \\
&= \varprojlim_n \mathrm{Hom}^{fact}_{W(\mathsf{k})}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), \mathbf{Z}_p) \otimes B/p^nB \\
&= \varprojlim_n \mathrm{Hom}^{fact}_{W(\mathsf{k})}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), B/p^nB) \\
&= \varprojlim_n \mathrm{Hom}_{W(\mathsf{k}),conts}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), B/p^nB) \\
&= \mathrm{Hom}_{W(\mathsf{k}),conts}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), B).
\end{aligned}$$

The second assertion is immediate from the definition of ϕ_f, and the third is clear, since any eigenform defines an algebra homomorphism

$$\mathbf{T}_0(W(\mathsf{k}),\mathrm{N}) \longrightarrow B$$

(its system of eigenvalues). □

In most situations, it will be sufficient with the case $W(\mathsf{k}) = \mathbf{Z}_p$, which is the most general. In the theory of deformations of Galois representations, however, we will want to base-change to $W(\mathsf{k})$.

In particular, the Corollary applies to classical cuspforms. Hence, for *any* k and ν, a classical cuspform f of level $\mathrm{N}p^\nu$, weight k, and defined over B, determines a continuous homomorphism ϕ_f of $\mathbf{Z}_p$-modules $\mathbf{T}_0 \longrightarrow B$. If f is a normalized eigenform for $\mathbf{T}_0$ (hence in particular has a nebentypus), then ϕ_f is an algebra homomorphism whose restriction to $G(\mathrm{N})$ is the character determined by the weight and nebentypus[1].

Remark: The reference to the various topologies in the statement of this general duality result should be carefully noted. A parabolic generalized p-adic modular function defined over a p-adic ring B may be identified with a continuous $\mathbf{Z}_p$-module homomorphism $\mathbf{T}_0 \longrightarrow B$, provided one gives $\mathbf{T}_0$ its inverse limit topology (of which we give a more intrinsic description below) and B its p-adic topology. For example, the identity map $\mathbf{T}_0 \xrightarrow{id} \mathbf{T}_0$ does *not* correspond to a modular function (because the p-adic topology is strictly finer than the inverse limit topology); in other words, there is no "universal" parabolic modular function. This leads to the definition of a family of modular functions in the next section.

Since the inverse limit topology on $\mathbf{T}_0$ appears in such a central manner in the discussion above, one would like to be able to give an intrinsic characterization of it, independent of the particular description of $\mathbf{T}_0$ as an inverse limit of finite $\mathbf{Z}_p$-algebras (of which there are many, since one may work with level $\mathrm{N}p^\nu$ for any $\nu \geq 1$, and see also section III.3). This is indeed possible, at least when $B = \mathcal{O}_K$ for some finite extension $K/\mathbf{Q}_p$.

Proposition III.1.4 *Let $B = \mathcal{O}_K$ for some finite extension $K/\mathbf{Q}_p$. Under the isomorphism*

$$\mathbf{T}_0(B,\mathrm{N}) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{V}_{par}, B),$$

the inverse limit topology of $\mathbf{T}_0(B,\mathrm{N})$ corresponds to the compact-open topology on $\mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{V}_{par}, B)$ (where $\mathbf{V}_{par}$ and B are given their p-adic topologies).

[1] In the preceding chapter, we recalled Hida's theory of the *ordinary* part $e\mathbf{T}_0$ of the Hecke algebra, which is a direct summand of $\mathbf{T}_0$, and corresponds to the "unit-root eigenspace" for the U operator. It is immediate, then, that if the eigenform f satisfies $\mathrm{U}f = \lambda f$ with λ a p-adic unit, then ϕ_f factors through the ordinary algebra $e\mathbf{T}_0$.

Proof: The inverse limit topology is generated by open sets of the form

$$\{\mathrm{T} \in \mathbf{T}_0 \,|\, \mathrm{T}f \in \mathcal{U},\ \forall f \in \mathsf{S}^k\},$$

where $\mathcal{U}$ runs over the open subsets of B and k runs over the integers, which are open in the compact-open topology (because the S^k, being free of finite rank over $\mathbf{Z}_p$, are compact in the p-adic topology). Hence, it is clear that the compact-open topology is finer than the inverse limit topology.

For the converse, we want to show that sets of the form

$$\{\mathrm{T} \in \mathbf{T}_0 \,|\, \mathrm{T}(K) \subset \mathcal{U}\},$$

($K \subset \mathbf{V}_{par}$ compact, $\mathcal{U} \subset B$ open) are open in the inverse limit topology. It is clearly enough to consider the case where K is a compact submodule of $\mathbf{V}_{par}$ and $\mathcal{U} = p^n B$. Since K is compact in the p-adic topology and $\mathbf{Z}_p$-free, it must be $\mathbf{Z}_p$-free of finite rank; let $(f_1, \ldots, f_r)$ be a basis. Since S is dense in $\mathbf{V}_{par}$, one may choose $(g_1, \ldots, g_r)$ such that $g_i \in \mathsf{S}^k$ (for some fixed, sufficiently large k) and $g_i \equiv f_i \pmod{p^\nu}$ with $\nu > n$. Then it is clear that

$$\begin{aligned}\{\mathrm{T} \,|\, \mathrm{T}(K) \subset p^n B\} &= \{\mathrm{T} \,|\, \mathrm{T}f_i \in p^n B,\ i = 1, \ldots, r\} \\ &= \{\mathrm{T} \,|\, \mathrm{T}g_i \in p^n B,\ i = 1, \ldots, r\},\end{aligned}$$

which is clearly open in the inverse limit topology (it is the inverse image under the canonical projection of an open subset of $\mathfrak{h}_k$). Thus, the two topologies are equal. □

III.1.3 The non-parabolic case

To get analogous results for the full ring of p-adic modular functions (not necessarily parabolic), one must somehow get around the fact that the constants in $\mathbf{V}$ will necessarily pair to zero in the pairing (III.1). This turns out not to be too difficult, and since we will not use it later, we will only sketch the results.

Let K be a finite extension of $\mathbf{Q}_p$, and let $B = \mathcal{O}_K$. We have already seen that we have the perfect pairing

$$\mathcal{H}'_k(B, \mathrm{N}) \times m_k(B, \mathrm{N}) \longrightarrow B,$$

where m_k is defined as above, rather than a pairing for D'_k, since it is clear that a sum of modular forms over K (the field of fractions of B) that has integral q-expansion except for its a_0 term will define a map $\mathcal{H}'_k \longrightarrow B$. Note that

$$m_k(B, \mathrm{N}) \cong \{f \in \mathsf{D}'_k(K, \mathrm{N}) | f(q) \in K + qB[[q]]\},$$

so that going to $m_k(B, \mathrm{N})$ gives precisely the desired space.

It is again possible to pass to the limit situation. Since we have the obvious maps $m_k(B,\mathrm{N}) \hookrightarrow m_{k+1}(B,\mathrm{N})$, we may consider the limit, and define

$$m_\infty(B,\mathrm{N}) = \varinjlim_k m_k(B,\mathrm{N}).$$

Then note that

$$m_\infty(B,\mathrm{N}) = \varinjlim_k \frac{\mathsf{D}_k(B,\mathrm{N})+K}{K} = \frac{\mathsf{D}(B,\mathrm{N})+K}{K},$$

so that $m_\infty(B,\mathrm{N})$ is not a completely mysterious space. (It may also be described as in the preceding paragraph, as the space of sums of modular forms over K which have integral q-expansion except possibly for the first coefficient.)

As before, we get a pairing

$$\mathbf{T}(B,\mathrm{N}) \times m_\infty(B,\mathrm{N}) \longrightarrow B,$$

which, as before, gives a duality between $\mathbf{T}(B,\mathrm{N})$ with its inverse image topology and $\overline{m}_\infty(B,\mathrm{N})$, where the bar indicates the p-adic completion. (We have here used the fact that

$$\mathsf{D}' = \varinjlim_k \mathsf{D}'_k$$

is dense in $\mathbf{V}$, and hence that its Hecke algebra is the full Hecke algebra, as noted above.) The only difficulty, then, is to determine the relation between $\mathbf{V}$ and the p-adic completion of m_∞. The elements of $\overline{m_\infty}$ are limits of sequences $f_i \in m_\infty(B,\mathrm{N})$ of modular forms whose q-expansions are integral except perhaps for the a_0 term, and the p-adic norm is the p-adic norm on q-expansions shorn of their a_0 term, i.e., if $f = \sum a_n q^n$, with a_n p-integral for $n \geq 1$, then we set

$$\|f\| = \sup_{n\geq 1} |a_n|,$$

where $|\cdot|$ denotes the p-adic norm on B. Giving $m_\infty(B,\mathrm{N})$ this topology, we have a continuous surjective Hecke-equivariant map

$$\mathsf{D}(B,\mathrm{N}) + K \longrightarrow m_\infty(B,\mathrm{N}),$$

which extends by continuity to a map

$$\mathbf{V}(B,\mathrm{N}) + K \longrightarrow \overline{m}_\infty(B,\mathrm{N}).$$

We claim this is still onto. Let $[f]$ denote the image in m_∞ of $f \in \mathsf{D}+K$; it is clear that $\|[f]\| \leq \|f\|$, where of course we use the correct norm in each space. However, one can make a canonical choice of the lifting f by requiring $f(q) \in qB[[q]]$ (which, since the constants belong to K, can always be done!), in which case we will have $\|f\| = \|[f]\|$. Note also that such a lifting f will belong to D, and not only to $\mathsf{D}+K$. Thus, given a Cauchy sequence in m_∞, we can lift it to a Cauchy sequence in D, whose limit will be mapped to the limit of the given sequence, so that the map on completions is onto. The kernel consists precisely of the constants $K \subset \mathbf{V} + K$, so that we get an exact sequence of Hecke-equivariant $\mathbf{Z}_p$-algebra homomorphisms

$$0 \longrightarrow K \longrightarrow \mathbf{V}(B,\mathrm{N}) + K \longrightarrow \overline{m}_\infty(B,\mathrm{N}) \longrightarrow 0.$$

To summarize, we have shown:

Proposition III.1.5 *Let K be a finite extension of $\mathbf{Q}_p$, and let $B = \mathcal{O}_K$. Then the pairing*

$$\begin{array}{ccc} \mathbf{V}(B,\mathrm{N}) \times \mathbf{T}(B,\mathrm{N}) & \longrightarrow & B \\ (f,\mathrm{T}) & \mapsto & a_1(\mathrm{T}f) \end{array}$$

induces isomorphisms

$$\mathbf{T}(B,\mathrm{N}) \cong \mathrm{Hom}_B(\mathbf{V}_1(B,\mathrm{N}), B)$$

(continuous homomorphisms of B-modules) and

$$\mathbf{V}_1(B,\mathrm{N}) \cong \mathrm{Hom}_{B,conts}(\mathbf{T}(B,\mathrm{N}), B),$$

where we give $\mathbf{T}(B,\mathrm{N})$ the inverse limit topology, and where

$$\mathbf{V}_1(B,\mathrm{N}) = \frac{\mathbf{V}(B,\mathrm{N}) + K}{K}.$$

Thus a continuous B-module homomorphism $\mathbf{T}(B,\mathrm{N}) \longrightarrow B$ determines a unique element of $\mathbf{V}_1(B,\mathrm{N})$; however, in general there is no canonical Hecke-equivariant way to determine a lifting to $\mathbf{V}(B,\mathrm{N})$. If the map is a B-*algebra* homomorphism, the corresponding element of $\mathbf{V}_1(B,\mathrm{N})$ is an eigenform (up to constants), and hence has a well-defined weight. Provided the weight is not zero (i.e., provided $G(\mathrm{N})$ is not mapped to 1), we can then choose canonically a constant term in the q-expansion, so that such a map corresponds to a well-defined eigenform in $\mathbf{V}(B,\mathrm{N})[1/p]$ which has q-expansion $f(q) \in K + qB[[q]]$. The example of

$$f = \frac{1}{p}\mathrm{E}_{p-1}$$

shows that it is not possible, in general, to assume that the resulting eigenform is in $\mathbf{V}(B,\mathrm{N})$ itself.

Applying the preceding proposition to the case $B = \mathbf{Z}_p$ and using an inverse limit argument analogous to that in the proof of Corollary III.1.3, we get an analogous result. Recall that $\mathbf{T} = \mathbf{T}(\mathbf{Z}_p, \mathrm{N})$, and give $\mathbf{T}$ its inverse limit topology; then:

Corollary III.1.6 *For any p-adic ring B with the p-adic topology, we have*

$$\begin{array}{c} \mathbf{V}_1(B,\mathrm{N}) \cong \mathrm{Hom}_{\mathbf{Z}_p,conts}(\mathbf{T}, B) \\ f \mapsto \phi_f \end{array}$$

(continuous homomorphisms of $\mathbf{Z}_p$-modules), where

$$\phi_f(\mathrm{T}) = a_1(\mathrm{T}f)$$

and

$$\mathbf{V}_1(B,\mathrm{N}) = \frac{\mathbf{V}(B,\mathrm{N}) + K}{K}.$$

Moreover, ϕ_f is an algebra homomorphism if and only if $f \in \mathbf{V}_1(B,\mathrm{N})$ is a normalized simultaneous eigenform for the Hecke and diamond operators (up to constants). In this case, $f \in \mathbf{V}_1(B,\mathrm{N})$ may be canonically lifted to an eigenform $f \in \mathbf{V}(B,\mathrm{N})[1/p]$ whose q-expansion satisfies $f(q) \in K + qB[[q]]$, provided $\phi_f(G(\mathrm{N})) \neq 1$, i.e., provided f is not of weight zero.

In what follows, we will usually prefer to consider only the case of parabolic modular functions, since the duality theory is then much simpler, but will occasionally mention the general case.

III.2 Families of Modular Forms

As we have observed in the preceding section, the general duality between parabolic p-adic modular functions and the Hecke algebra $\mathbf{T}_0$ is complicated by the necessity of keeping the topologies involved straight. The point of this section is to introduce a concept that alleviates the problem somewhat. As before, everything makes sense (and is true) for p-adically complete algebras over the Witt ring of a finite field; we will consider only the case $\mathsf{k} = \mathbf{F}_p$ (so that $W(\mathsf{k}) = \mathbf{Z}_p$), and leave the extension to the reader.

The concept of an "analytic family of p-adic modular forms" was first introduced by Serre in [Se73]. He considered there the Iwasawa algebra $\Lambda = \mathbf{Z}_p[[\Gamma]]$ (the completed group ring with coefficients in $\mathbf{Z}_p$ of the pro-p-group $\Gamma = 1 + p\mathbf{Z}_p \subset \mathbf{Z}_p^\times$), given the inverse limit topology, with which it is a compact $\mathbf{Z}_p$-algebra. Recall that choosing a topological generator γ of Γ defines an isomorphism of topological $\mathbf{Z}_p$-algebras

$$\begin{array}{ccc} \mathbf{Z}_p[[\Gamma]] & \longrightarrow & \mathbf{Z}_p[[T]] \\ \langle\gamma\rangle & \longmapsto & 1+T, \end{array}$$

where $\mathbf{Z}_p[[T]]$ is given the (p,T)-adic topology, and where we use angular brackets to distinguish elements of the group Γ from elements of $\mathbf{Z}_p$ (so that $\langle\gamma\rangle - \gamma$ is a nonzero element of the group ring). Then Serre defined an analytic family of modular forms to be a formal q-expansion

$$F(q) = A_0 + A_1 q + A_2 q^2 + \dots$$

where $A_j \in \Lambda$, such that, for every $k \in \mathbf{Z}_p$ and for a fixed $i \in \mathbf{Z}/(p-1)\mathbf{Z}$, the image of $F(q)$ under the map $\Lambda \longrightarrow \mathbf{Z}_p$ defined by $\langle\gamma\rangle \mapsto \gamma^k$ is a p-adic modular form of weight (i,k), i.e., belongs to $\mathsf{M}(\mathbf{Z}_p, \chi_{(i,k)}, \mathrm{N}; 1)$. To get a family of cusp forms, of course, we would require $A_0 = 0$ and that each specialization be a cusp form.

To translate this definition to our situation (in the cuspidal case), note that $\mathbf{T}_0$ is naturally an algebra over the algebra $\mathbf{\Lambda} = \mathbf{Z}_p[[\mathbf{Z}_p^\times]]$ topologically generated by the diamond operators $\langle x, 1\rangle$ for $x \in \mathbf{Z}_p$. Note that $\mathbf{\Lambda} = \Lambda[\mathbf{Z}/(p-1)\mathbf{Z}]$ is just a group ring over Λ, which (because $p \not| \, (p-1)$) we may decompose according to the powers of the Teichmüller character (which are the characters of $\mathbf{Z}/(p-1)\mathbf{Z}$); we write

$$\mathbf{\Lambda} = \bigoplus_{i \bmod (p-1)} \Lambda_{(i)}.$$

Then it is clear that an analytic family of p-adic modular forms in the sense above is a map $\mathbf{T}_0 \longrightarrow \Lambda_{(i)}$ which, when composed with the canonical maps $\Lambda \longrightarrow \mathbf{Z}_p$ given by $\langle\gamma\rangle \mapsto \gamma^k$, gives continuous maps $\mathbf{T}_0 \longrightarrow \mathbf{Z}_p$. This amounts to a map $\mathbf{T}_0 \longrightarrow \Lambda_{(i)}$ which is continuous when we give Λ its *inverse limit* topology. Thus, an analytic family of p-adic modular forms is *not* necessarily a generalized p-adic modular function defined over Λ, though it defines p-adic modular forms by "specialization to weight k" for every k. (The distinction may be understood as follows: a generalized p-adic modular function defined over Λ can be evaluated at any trivialized elliptic curve defined over *any* p-adically complete Λ-algebra; a map $\mathbf{T}_0 \longrightarrow \Lambda$ does not determine such a rule. Consider, for example, a trivialized curve defined over $\Lambda \otimes \mathbf{F}_p \cong \mathbf{F}_p[[T]]$.)

Given this re-interpretation of Serre's definition, it is clear that we may extend it as follows.

Definition III.2.1 *Let* $\mathbf{B}$ *be a p-adically complete topological* $\mathbf{Z}_p$*-algebra. A* $\mathbf{B}$*-valued family of parabolic p-adic modular functions is a continuous* $\mathbf{Z}_p$*-module homomorphism* $\mathbf{f} : \mathbf{T}_0 \longrightarrow \mathbf{B}$*. Given any continuous map* $\phi : \mathbf{B} \longrightarrow B$ *to a p-adic ring* B *(where we give* B *the p-adic topology), we denote by* $\mathbf{f}_\phi$ *the modular function defined over* B *corresponding to the composite homomorphism* $\phi \circ \mathbf{f}$*; we call* $\mathbf{f}_\phi$ *the "specialization via* ϕ*" (sometimes "to weight* ϕ*") of the family* $\mathbf{f}$*. Finally, we say* $\mathbf{f}$ *is a family of eigenforms if every specialization* $\mathbf{f}_\phi$ *is a simultaneous eigenfunction for the Hecke, diamond, and* U *operators (and hence for the action of* $\mathbf{T}_0$*).*

Remarks:

i. Of course, if the map $\mathbf{f}$ is continuous when we give $\mathbf{B}$ the p-adic topology, then we simply obtain a parabolic modular function defined over $\mathbf{B}$. This makes $\mathbf{B}$-valued families of modular functions a generalization of modular functions defined over $\mathbf{B}$.

ii. In many cases, the family $\mathbf{f}$ is determined by the set of all its specializations (for example, when $\mathbf{B}=\mathbf{\Lambda}$); when that is true, we will sometimes confuse the family $\mathbf{f}$ with the set $\{\mathbf{f}_\chi : \chi \in \mathrm{Hom}_{conts}(\mathbf{Z}_p^\times, \mathbf{Z}_p^\times)\}$ of its specializations.

iii. A family of parabolic modular functions has a q-expansion; simply define

$$\mathbf{f}(q) = \sum_{n\geq 1} \mathbf{f}(\mathrm{T}_n) q^n.$$

iv. The "universal family" of parabolic modular functions is simply the identity map $\mathbf{T}_0 \longrightarrow \mathbf{T}_0$.

The most interesting case of the above will be when $\mathbf{B}$ is a $\mathbf{\Lambda}$-algebra, in which case one can consider, for each $\chi \in \mathrm{Hom}_{conts}(\mathbf{Z}_p^\times, \mathbf{Z}_p^\times)$, the map

$$\phi_\chi : \mathbf{B} \longrightarrow B_\chi = \mathbf{B}/P_\chi\mathbf{B},$$

where $P_\chi\mathbf{B}$ denotes the ideal generated by the elements $\langle x, 1\rangle - \chi(x) \in \mathbf{\Lambda}$. If these maps are continuous when we give the quotient the p-adic topology, we call them "specialization to weight χ", and write $\mathbf{f}_\chi$ for the corresponding specializations of a family of modular forms $\mathbf{f}$. Note that we then have $\mathbf{f}_\chi \in \mathsf{M}(B_\chi, \chi, \mathrm{N}; 1)$. For example, if $\mathbf{B} = \mathbf{\Lambda}$, we have $B_\chi = \mathbf{Z}_p$ for all χ, and we recover Serre's situation.

Remark: There is one important aspect in which the above definition is weaker than Serre's: the restriction to parabolic modular functions. In fact, the best-known example of an analytic family, Serre's Eisenstein family G_χ, is not parabolic. The problem, of course, is the fact that the duality between modular functions and Hecke operators gets complicated in the non-parabolic case. In particular, it is not clear that a family of eigenforms $\mathbf{f} : \mathbf{T} \longrightarrow \Lambda$ in our sense determines a family of eigenforms in Serre's sense, i.e., that there is a well-defined $A_0 \in \Lambda$ which is the constant term in the q-expansion of $\mathbf{f}$. In fact, this is not always true; given the duality theorems, the most we can expect is an element of the fraction field of Λ. Serre has shown (in [Se73]) that we have two cases. First, suppose i is not divisible by $p-1$. Then, given modular functions F_k of p-adic weight (i, k), they are the specializations of an analytic family of modular forms ("with values in Λ") if and only if there are elements $A_i \in \Lambda$, $i \geq 1$, such that A_i specializes (via $\gamma \mapsto \gamma^k$) to the i^{th} coefficient of the q-expansion of F_k for each k. (In other words, if the coefficients $a_1(F_k), a_2(F_k), \ldots$ are specializations of elements of Λ, then so is the first coefficient $a_0(F_k)$.) This avoids the difficulties with the integrality of the zero-th coefficient, and shows that in this case an analytic family of (non-parabolic) p-adic modular forms with values in Λ *is* a continuous map $\mathbf{T} \longrightarrow \Lambda$. In the case when $i \equiv 0 \pmod{p-1}$, however, the Eisenstein family already shows that this cannot be the case; in this case Serre shows that if $a_1(F_k), a_2(F_k), \ldots$ are specializations of elements of Λ, then $a_0(F_k)$ will be a specialization of an element of the fraction field of Λ of the form

$$\frac{c}{\gamma^r - 1},$$

with $c \in \Lambda$. In any case, we may identify "analytic families of modular forms with values in Λ" with maps $\mathbf{T} \longrightarrow \Lambda$, provided we allow the zero-th coefficient in the "analytic family" to belong to the fraction field of Λ. This extends to families with values in $\mathbf{\Lambda}$, for trivial reasons, but it is not clear what happens for more general rings.

We should note the following obvious lemma:

Lemma III.2.2 *Let* B *be a* $\mathbf{\Lambda}$*-algebra, and assume that the intersection of all the ideals* $P_\chi\mathrm{B}$ *of* B *is zero. Then a family of parabolic modular forms* $\mathbf{f} : \mathbf{T}_0 \longrightarrow \mathrm{B}$ *with values in* B *is a family of eigenforms if and only if the map* $\mathbf{f}$ *is a continuous* $\mathbf{Z}_p$-algebra *homomorphism.*

Proof: clear. □

Analytic families with values in Λ can be detected, as Serre showed in [Se73], by congruence properties. Since $\mathbf{\Lambda}$ is just a sum of copies of Λ, the same holds for families with values in $\mathbf{\Lambda}$. We refer to Serre's paper for the precise statements: see [Se73, Section 4.4].

Finally, we note that one can use Serre's Eisenstein family to construct any number of examples of $\mathbf{\Lambda}$-valued families of modular forms. For this, let f be a cusp form of p-adic weight (i, k_0), and let E_k, with $k \in \mathbf{Z}_p$ denote (the specializations of) Serre's analytic family of Eisenstein series of weight $(0, k)$. Recall that $\mathrm{E}_0 = 1$. Then it is clear (say, from congruence properties) that setting $\mathbf{f}_k = f \cdot \mathrm{E}_{k-k_0}$ defines a family of modular forms with values in Λ, whose specialization to weight k_0 is precisely f. We describe this last fact by saying that $\mathbf{f}$ is a "spreading-out" (over Λ) of the modular form f. Thus, this construction not only provides examples of families of modular forms but also shows that one can spread out any modular form to an analytic family containing it. Unfortunately, the resulting family has few good properties, because the specializations of $\mathbf{f}$ are not eigenforms for the Hecke algebra, except in very special cases. (For example, in the "rank one" case of Hida's theory of the ordinary part, we can get a family of eigenforms by taking the ordinary part of $\mathbf{f}$.) We will later show one way of obtaining a family of eigenforms by twisting a classical modular form by wild characters. Finding other methods for generating families of eigenforms would have significant corollaries for the deformation theory of Galois representations discussed ahead, and is an important open problem.

III.3 Changing the Level

As we have seen above, the ring $\mathbf{V}$ of generalized p-adic modular functions may be realized as the closure of a union of $\mathbf{Z}_p$-modules of finite rank, by way of divided congruences. In this section, we give another description of $\mathbf{V}$ in the same spirit, this time taking classical modular forms of constant weight but varying level. The advantage of this description is that it does not necessitate the consideration of congruences of classical forms, so that theorems about classical modular forms over $\mathbf{Z}_p$ can be more easily applied to the situation.

Let us fix a weight $k \geq 2$. We begin by recalling (see Section I.3.5) that we have defined canonical inclusions

$$M(B, k, \mathrm{N}p^\nu) \hookrightarrow \mathbf{V}(B, \mathrm{N}),$$

and by noting that these submodules of $\mathbf{V}$ form a direct system via the canonical inclusions

$$M(B,k,\mathrm{N}p^{\nu}) \hookrightarrow M(B,k,\mathrm{N}p^{\nu+1}).$$

(Note that all of these maps preserve q-expansions.) Thus, if we set

$$M(B,k,\mathrm{N}p^{\infty}) = \varinjlim_{\nu} M(B,k,\mathrm{N}p^{\nu}),$$

we get a q-expansion preserving inclusion

$$M(B,k,\mathrm{N}p^{\infty}) \hookrightarrow \mathbf{V}(B,\mathrm{N}).$$

Similarly, by restricting to cusp forms, we may define $S(B,k,\mathrm{N}p^{\infty})$ and get a canonical inclusion preserving q-expansions,

$$S(B,k,\mathrm{N}p^{\infty}) \hookrightarrow \mathbf{V}_{par}(B,\mathrm{N}).$$

We are interested in considering the closure of the image (which is, by the q-expansion principle, the p-adic completion of $S(B,k,\mathrm{N}p^{\infty})$); let us denote it by $\overline{S}(B,k,\mathrm{N}p^{\infty})$. It is clear that the proofs given in Section III.1 allow us to show:

Proposition III.3.1 *Let B be a p-adic ring, and let*

$$\mathrm{h}_k(B,\mathrm{N}p^{\infty}) = \varprojlim_{\nu} \mathrm{h}(B,k,\mathrm{N}p^{\nu}),$$

where $\mathrm{h}(B,k,\mathrm{N}p^{\nu})$ denotes the Hecke algebra corresponding to the space $S(B,k,\mathrm{N}p^{\nu})$ of cusp forms of weight k and level $\mathrm{N}p^{\nu}$. Let $\mathrm{h}_k(B,\mathrm{N}p^{\infty})$ the inverse limit topology induced by the p-adic topology on the $\mathrm{h}(B,k,\mathrm{N}p^{\nu})$. Then the map $f \mapsto \phi_f$ defined as above gives an isomorphism

$$\overline{S}(B,k,\mathrm{N}p^{\infty}) \xrightarrow{\;\sim\;} \mathrm{Hom}_{\mathbf{Z}_p,conts}(\mathrm{h}_k(\mathbf{Z}_p,\mathrm{N}p^{\infty}),B).$$

Proof: Clear; note that the assumption $k \geq 2$ avoids all questions about base change. □

In fact, we have shown nothing new, because of the following result, which was stated without proof in [Hi86a], and whose proof is to appear in a forthcoming paper.

Theorem III.3.2 [**Hida**] *Let $k \geq 2$ and let B be p-adic ring. Then the image of $S(B,k,\mathrm{N}p^{\infty})$ under the canonical inclusion is dense in $\mathbf{V}_{par}(B,\mathrm{N})$. Equivalently, the surjection*

$$\mathbf{T}_0(B,\mathrm{N}) \twoheadrightarrow \mathrm{h}_k(B,\mathrm{N}p^{\infty})$$

defined by the canonical inclusion is an isomorphism.

Proof: The equivalence of the two statements follows from (III.1.3) and the preceding proposition (and one may even restrict to the case $B=\mathbf{Z}_p$).

Step 1: We note, first, that Hida has shown ([Hi86a]) that there are Hecke-invariant inclusions

$$M(B,k,\mathrm{N}p^\infty)\otimes B/p^nB \hookrightarrow M(B,k+1,\mathrm{N}p^\infty)\otimes B/p^nB,$$

given by multiplication by the appropriate Eisenstein series. These induce Hecke-invariant inclusions

$$\overline{S}(B,k,\mathrm{N}p^\infty) \hookrightarrow \overline{S}(B,k+1,\mathrm{N}p^\infty).$$

Since the Hecke actions extend naturally by continuity to the p-adic completions we get surjections

$$\mathfrak{h}_{k+1}(B,\mathrm{N}p^\infty) \longrightarrow\!\!\!\!\!\rightarrow \mathfrak{h}_k(B,\mathrm{N}p^\infty).$$

Thus we get a diagram of surjections,

$$\begin{tikzcd}
 & \mathfrak{h}_{k+1}(B,\mathrm{N}p^\infty) \arrow[d] \\
 & \mathfrak{h}_k(B,\mathrm{N}p^\infty) \arrow[d] \\
\mathbf{T}_0(B,\mathrm{N}) \arrow[uur] \arrow[ur] \arrow[dr] & \vdots \arrow[d] \\
 & \mathfrak{h}_2(B,\mathrm{N}p^\infty)
\end{tikzcd}$$

and it follows that it is enough to prove that the map

$$\mathbf{T}_0(B,\mathrm{N}) \longrightarrow\!\!\!\!\!\rightarrow \mathfrak{h}_2(B,\mathrm{N}p^\infty)$$

is an isomorphism, or equivalently, that the image of $S(B,2,\mathrm{N}p^\infty)$ is dense in $\mathbf{V}_{par}$.

Step 2: The crucial step in the proof is the following result of Hida ([Hi]):

Theorem III.3.3 **[Hida]** *The maps* $\mathfrak{h}_k(B,\mathrm{N}p^\infty) \longrightarrow \mathfrak{h}_{k-1}(B,\mathrm{N}p^\infty)$ *are all isomorphisms.*

Hida proves this by studying the representation of the Hecke algebras in question on the parabolic cohomology of congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$, and by invoking a result of Shimura and Ohta. (The result is stated by Shimura in [Shi68] and proved, in the quaternionic case, by Ohta in [Oht82]; using recent results of Harder, one sees that Ohta's proof applies without change in our situation.) We refer the reader to Hida's forthcoming paper.

Step 3: The full result now follows easily. Note, first, that Hida's theorem implies that the maps $\overline{S}_k(B,\mathrm{N}p^\infty) \longrightarrow \overline{S}_{k+1}(B,\mathrm{N}p^\infty)$ are isomorphisms. Reducing mod p^n, we get that the maps

$$S_k(B/p^nB,\mathrm{N}p^\infty) \longrightarrow S_{k+1}(B/p^nB,\mathrm{N}p^\infty)$$

are isomorphisms, so that we have the following:

Corollary III.3.4 *Let $n \geq 1$. Fix $k_0 \in \mathbf{Z}$, $k_0 \geq 2$. Then, for any $k \geq 2$ and any classical modular form f of weight k and level $\mathrm{N}p^{\nu}$, there exists $\mu \geq \nu$ and a classical modular form g of weight k_0 and level $\mathrm{N}p^{\mu}$ such that $f \equiv g \pmod{p^n}$.*

Now, to conclude the proof, we need to show that the map

$$\mathbf{T}_0 \longrightarrow \mathrm{h}_k(B, \mathrm{N}p^{\infty})$$

is an isomorphism (for any fixed k). Since we know it is surjective, we need only show that it is injective. To see this, for each

$$\chi \in \mathrm{Hom}_{conts}(\mathbf{Z}_p^{\times}, \mathbf{Z}_p^{\times}),$$

let P_{χ} be the ideal of $\mathbf{\Lambda} = \mathbf{Z}_p[[\mathbf{Z}_p^{\times}]]$ generated by

$$\{\langle x, 1\rangle - \chi(x) : x \in \mathbf{Z}_p^{\times}\}.$$

Then it is clear from the duality theorems that $\mathbf{T}_0/P_{\chi}\mathbf{T}_0$ is dual to the space $\mathsf{S}(B, \chi, \mathrm{N}; 1)$ of p-adic cusp forms of weight χ (because this space is precisely the subspace of $\mathbf{V}_{par}$ consisting of forms satisfying $\langle x, 1\rangle f = \chi(x) f$). We want to look at the map

$$\mathbf{T}_0/P_{\chi}\mathbf{T}_0 \longrightarrow \mathrm{h}_k(B, \mathrm{N}p^{\infty})/P_{\chi}\mathrm{h}_k(B, \mathrm{N}p^{\infty}).$$

Dualizing, we get the inclusion

$$\overline{S}_k(B, \mathrm{N}p^{\infty}) \cap \mathsf{S}(B, \chi, \mathrm{N}; 1) \hookrightarrow \mathsf{S}(B, \chi, \mathrm{N}; 1).$$

We claim that this is a surjection. Given $f \in \mathsf{S}(B, \chi, \mathrm{N}; 1)$, we can find a sequence of classical cusp forms $f_n \in S(B, k_n, \mathrm{N})$ such that

$$f(q) \equiv f_n(q) \pmod{p^n},$$

so that $f_n \to f$ in the q-expansion topology. By the corollary, we can find $g_n \in S(B, k, \mathrm{N}p^{\nu(n)})$ such that $g_n(q) \equiv f_n(q) \pmod{p^n}$, and then we get that $g_n \to f$, so that $f \in \overline{S}_k(B, \mathrm{N}p^{\infty})$, proving our claim.

Dualizing again, we see that the map

$$\mathbf{T}_0/P_{\chi}\mathbf{T}_0 \longrightarrow \mathrm{h}_k(B, \mathrm{N}p^{\infty})/P_{\chi}\mathrm{h}_k(B, \mathrm{N}p^{\infty})$$

is an isomorphism (for any χ). It follows that the map

$$\mathbf{T}_0 \longrightarrow \mathrm{h}_k(B, \mathrm{N}p^{\infty})$$

is injective: if T is in the kernel, then $\mathrm{T} \in P_{\chi}\mathbf{T}_0$ for all χ; since the divided congruences of p-adic cusp forms are dense in $\mathbf{V}_{par}$, we have $\bigcap P_{\chi}\mathbf{T}_0 = 0$, so that $T = 0$. This proves the theorem. □

III.4 Deformations of Residual Eigenforms

In [Ma], Mazur has considered the problem of studying the deformations, in a sense to be defined below, of representations of the absolute Galois group of the field of rational numbers in $GL_2(\mathsf{k})$, where k is a finite field, which he calls a "residual representation". It is well known that some representations of this kind may be obtained from modular forms, and it is an interesting question whether it is true that the deformations of a residual representation arising from a modular form are in some sense modular. In this and the following section, we consider *modular* deformations of a residual representation, i.e., those arising from modular forms, and show that one can construct a universal modular representation which corresponds to a family of p-adic modular functions which are eigenforms for the Hecke algebra. We begin, in this section, by considering deformations of "residual eigenforms", i.e., p-adic modular functions defined over a finite field k which are eigenforms for the Hecke algebra. For technical reasons, we restrict the discussion to parabolic modular functions.

Definition III.4.1 *Let B be a p-adic ring. A* Katz eigenform defined over B *is a parabolic generalized modular function which is an eigenform for the Hecke algebra and which is* normalized *in the sense that the coefficient of q in $f(q)$ is equal to 1; we denote the corresponding homomorphism $\mathbf{T}_0 \longrightarrow B$ by φ_f. This homomorphism induces characters $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ and $\varepsilon : (\mathbf{Z}/\mathrm{N}\mathbf{Z})^\times \longrightarrow B^\times$, which we call, respectively, the* weight *and the* nebentypus *of f.*

Note that the weight and nebentypus of an eigenform as defined here coincide with the classical weight and nebentypus for classical eigenforms of level N, but *not* for classical eigenforms of level $\mathrm{N}p^\nu$; in the latter case, ε is the "prime to p" part of the nebentypus character. If we need to refer to both terminologies, we will call what we have just defined the "p-adic weight" and "p-adic nebentypus".

Definition III.4.2 *A* residual eigenform *is a Katz eigenform $\bar{f}$ defined over a finite field k. Its weight is then necessarily a power of the "Teichmüller" character $\mathbf{Z}_p^\times \longrightarrow \mathbf{F}_p^\times \hookrightarrow \mathsf{k}^\times$ (i.e., the character induced by the canonical map $\mathbf{Z}_p \longrightarrow \mathbf{F}_p$).*

If $\bar{f}$ is a residual eigenform, it follows from Proposition I.3.3 that we have $\bar{f} \in \mathbf{V}_{1,1}$, and hence (because it has a weight) that $\bar{f}$ is the reduction of a classical modular form of level N defined over $W(\mathsf{k})$ (the Witt ring of k) — see [Ka75a]. It then follows from a lemma of Serre and Deligne (see ahead) that $\bar{f}$ is in fact the reduction of a classical *eigenform* of level N defined over some (totally ramified) extension of $W(\mathsf{k})$.

Definition III.4.3 *A* deformation *of a residual eigenform $\bar{f}$ defined over a finite field k is a Katz eigenform f defined over an artinian local $W(\mathsf{k})$-algebra A with residue field k such that $f \mapsto \bar{f}$ under the residue map. More generally, one may take A to be a complete noetherian local $W(\mathsf{k})$-algebra with residue field k whose quotients by powers of the maximal ideal are artinian.*

Typically, we will be looking at deformations to $W(\mathsf{k})$-algebras A which are either finite extensions of $W(\mathsf{k})$ or their quotients, and at families of such. We will later need to modify this notion somewhat, in order to adjust it to the situation of deformations of Galois representations.

III.4.1 Universal deformations

Fix a residual eigenform $\bar{f} \in \mathbf{V}(\mathsf{k}, \mathrm{N})$ over a finite field k. Let $\mathcal{C}$ denote the category of artinian local $\mathbf{Z}_p$-algebras A with residue field k, and let $\hat{\mathcal{C}}$ denote the category of complete noetherian local $W(\mathsf{k})$-algebras B with residue field k such that B/m^n is an object of $\mathcal{C}$ for any power of the maximal ideal $m \subset B$.

Consider the functor $\mathbf{F} : \mathcal{C} \longrightarrow \mathbf{Set}$ defined by

$$\begin{aligned} \mathbf{F}(A) &= \{\text{deformations of } \bar{f} \text{ defined over } A\} \\ &= \{\varphi \in \mathrm{Hom}_{\mathbf{Z}_p-alg}(\mathbf{T}_0, A) \mid \bar{\varphi} = \varphi_{\bar{f}}\} \\ &= \{\varphi \in \mathrm{Hom}_{W(\mathsf{k})-alg}(\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}), A) \mid \bar{\varphi} = \varphi_{\bar{f}}\} \end{aligned}$$

where $\mathrm{Hom}_{W(\mathsf{k})-alg}$ denotes continuous homomorphisms of $W(\mathsf{k})$-algebras (note: A is given the p-adic topology), and $\bar{\varphi}$ denotes the composition of φ with the residue map $A \longrightarrow \mathsf{k}$. We want to study the representability of this functor.

Since $\bar{f}$ is an eigenform, it corresponds to a continuous algebra homomorphism $\bar{\varphi} : \mathbf{T}_0 \longrightarrow \mathsf{k}$; we denote the kernel of this homomorphism by $\mathfrak{m}$. Let R denote the localization of $\mathbf{T}_0$ at $\mathfrak{m}$, with the induced topology. Then every continuous $\varphi : \mathbf{T}_0 \longrightarrow A$ lifting $\bar{\varphi}$ factors through the canonical map $\mathbf{T}_0 \longrightarrow \mathrm{R}$. This suggests that R will represent our functor; the first difficulty is that R is not complete, so that we need to pass to its completion, which we denote by $\hat{\mathrm{R}}$. Then, of course, for any p-adic ring B (complete local noetherian with residue field k), given a continuous algebra homomorphism $\mathbf{T}_0 \longrightarrow B$ lifting $\bar{\varphi}$, we get a continuous algebra homomorphism $\mathrm{R} \longrightarrow B$, which then extends to the completion, giving a continuous algebra homomorphism $\hat{\mathrm{R}} \longrightarrow B$. The converse, however, is not true (unless $\mathsf{k} = \mathbf{F}_p$), because any conjugate of $\bar{f}$ under the Galois group $\mathrm{Gal}(\mathsf{k}/\mathbf{F}_p)$ will determine the same maximal ideal $\mathfrak{m}$, and hence the same completion $\hat{\mathrm{R}}$. In other words, $\hat{\mathrm{R}}$ actually represents the functor

$$\mathbf{G}(A) = \bigcup_{\sigma \in \mathrm{Gal}(\mathsf{k}/\mathbf{F}_p)} \{\text{deformations of } \bar{f}^{\sigma} \text{ defined over } A\}.$$

We prefer to separate the several Galois conjugates; this can be done by base-changing to the Witt ring $W(\mathsf{k})$. Thus, let $\mathfrak{m}$ now denote the kernel of the map $\mathbf{T}_0(W(\mathsf{k}), \mathrm{N}) \longrightarrow \mathsf{k}$ corresponding to $\bar{f}$, and let $\mathrm{R}(\bar{f})$ denote the completion of $\mathbf{T}_0(W(\mathsf{k}), \mathrm{N})$ at $\mathfrak{m}$. Then clearly the deformations of $\bar{f}$ to B correspond precisely to maps $\mathrm{R}(\bar{f}) \longrightarrow B$, and we get:

Proposition III.4.4 *For any residual parabolic eigenform $\bar{f} \in \mathbf{V}(\mathsf{k},\mathrm{N})$, there exists a complete local $W(\mathsf{k})$-algebra $\mathrm{R}(\bar{f})$ and a family of parabolic modular functions $\mathbf{f}$: $\mathbf{T}_0(W(\mathsf{k}),\mathrm{N}) \longrightarrow \mathrm{R}(\bar{f})$ which is universal for the deformations of $\bar{f}$; that is, so that for every deformation f defined over a p-adic ring B, there exists a continuous homomorphism $\alpha : \mathrm{R}(\bar{f}) \longrightarrow B$ such that f is the specialization of $\mathbf{f}$ via α.*

Remark: In general, the universal family $\mathbf{f}$ is *not* a modular function defined over $\mathrm{R}(\bar{f})$ (because the map $\mathbf{T}_0(W(\mathsf{k}),\mathrm{N}) \longrightarrow \mathrm{R}(\bar{f})$ is not continuous if we give $\mathrm{R}(\bar{f})$ the p-adic topology), so that the functor is not, in the strict sense, representable, in the sense that there is no "universal deformation" (to be precise, the universal deformation is not a modular form, but rather a family of modular forms).

It is interesting to ask if $\mathrm{R}(\bar{f})$ is noetherian. This is equivalent to asking if the "tangent space[2]" of $\mathrm{R}(\bar{f})$ is finite-dimensional, i.e., that

$$\mathbf{F}(\mathsf{k}[\epsilon]) = \mathrm{Hom}_{algebras}(\mathrm{R}(\bar{f}), \mathsf{k}[\epsilon])$$

is a finite-dimensional k-vector space, where $\mathsf{k}[\epsilon] = \mathsf{k} + \mathsf{k}\epsilon$ with $\epsilon^2 = 0$. For example, if $\mathrm{U}\bar{f} = 0$, one can obtain a (not very interesting) deformation to $\mathsf{k}[\epsilon]$ by $g = \bar{f} + \epsilon\,\mathrm{Frob}\,\bar{f}$; this satisfies $\mathrm{U}g = \epsilon g$ and has the same eigenvalues as $\bar{f}$ for the rest of the Hecke algebra. This shows that the dimension of the tangent space is always at least one. Is it finite? We have not been able to answer this in the general situation.

III.4.2 Deformations outside $\mathrm{N}p$

In the next section, we will relate deformations of residual eigenforms and deformations of residual Galois representations. From this point of view, the deformations considered above are not the right thing to consider. The problem is that the representation attached to a modular form depends only on its eigenvalues for the T_ℓ with $\ell \nmid \mathrm{N}p$, i.e., on the eigenvalues for the action of the *restricted* Hecke algebra $\mathbf{T}_0^\star$. Thus, if two Katz eigenforms have the same q-expansion coefficients a_n whenever n is prime to $\mathrm{N}p$, they will determine the same representation. This shows that we should look for deformations up to a weaker notion of equality.

Definition III.4.5 *Let B be a p-adic ring. We say two Katz eigenforms f and g defined over B are equal outside $\mathrm{N}p$ if we have $a_n(f) = a_n(g)$ whenever $(n, \mathrm{N}p) = 1$, where $a_n(h)$ denotes the coefficient of q^n in the q-expansion of h.*

Since Katz eigenforms are by definition normalized, this is equivalent to requiring that the eigenvalues under T_ℓ be the same for ℓ such that $(\ell, \mathrm{N}p) = 1$.

Definition III.4.6 *Given a residual eigenform $\bar{f}$, we say that a Katz eigenform g defined over B is a deformation of $\bar{f}$ outside $\mathrm{N}p$ if $a_n(g)$ reduces to $a_n(\bar{f})$ whenever $(n, \mathrm{N}p) = 1$.*

[2]In the sense of Schlessinger in [Sch68]; this is actually the "reduced Zariski tangent space", that is, the Zariski tangent space of $\mathrm{R}(\bar{f})/p\mathrm{R}(\bar{f})$.

We may reinterpret these definitions in terms of our duality theorems: we have $\mathbf{T}_0^\star \subset \mathbf{T}_0$, and we say that two algebra homomorphisms $\mathbf{T}_0 \longrightarrow B$ are equal outside $\mathrm{N}p$ if their restrictions to the subalgebra $\mathbf{T}_0^\star$ coincide. Then, of course, two Katz eigenforms are equal outside $\mathrm{N}p$ if and only if the corresponding algebra homomorphisms are. Similarly, if we denote the homomorphism corresponding to $\bar{f}$ by $\overline{\varphi} : \mathbf{T}_0 \longrightarrow \mathsf{k}$ and its restriction to $\mathbf{T}_0^\star$ by $\overline{\varphi}^\star$, we say a homomorphism $\psi : \mathbf{T}_0 \longrightarrow B$ is a deformation of $\overline{\varphi}$ outside $\mathrm{N}p$ if its restriction to $\mathbf{T}_0^\star$ reduces to $\overline{\varphi}^\star$.

Definition III.4.7 *We say a parabolic p-adic modular function $f \in \mathbf{V}$ is an eigenform outside* $\mathrm{N}p$ *if it is an eigenform for the action of the restricted Hecke algebra $\mathbf{T}_0^\star$. If f is normalized by $a_1(f) = 1$ and if we denote the continuous map associated to f by $\phi : \mathbf{T}_0 \longrightarrow B$, this is equivalent to requiring that the restriction $\phi^\star$ of ϕ to $\mathbf{T}_0^\star$ be an algebra homomorphism*

If we wish to consider deformations of a residual eigenform $\bar{f}$ outside $\mathrm{N}p$, we must of course identify all the maps $\mathbf{T}_0 \longrightarrow B$ which have the same restriction to $\mathbf{T}_0^\star$. This just amounts to considering algebra homomorphisms $\mathbf{T}_0^\star \longrightarrow B$, since it is easy to see that any such may be extended to a $\mathbf{Z}_p$-module homomorphism $\mathbf{T}_0 \longrightarrow B$, by considering the classical case [3].

Let us then consider deformations of $\bar{f}$ as an eigenform outside $\mathrm{N}p$, i.e., modular functions f reducing to $\bar{f}$ outside $\mathrm{N}p$ which are eigenforms under $\mathbf{T}_0^\star$, up to equality outside $\mathrm{N}p$. Equivalently, the question is to describe the continuous algebra homomorphisms $\mathbf{T}_0^\star(W(\mathsf{k}),\mathrm{N}) \longrightarrow B$ lifting the restriction $\overline{\varphi^\star}$ of $\overline{\varphi}$ to $\mathbf{T}_0^\star(W(\mathsf{k}),\mathrm{N})$. Let $\mathbf{m}^\star$ be the kernel of $\overline{\varphi}^\star$. It is clear that this situation is completely analogous to the preceding one, so that the functor F defined by

$$F(B) = \{\text{homomorphisms } \mathbf{T}_0^\star \longrightarrow B \text{ lifting } \overline{\varphi}^\star\}$$

is represented by the completion of $\mathbf{T}_0^\star(W(\mathsf{k}),\mathrm{N})$ at the ideal $\mathbf{m}^\star = \ker(\overline{\varphi}^\star)$. We will denote this ring by $\mathbf{R} = \mathbf{R}(\bar{f})$. Note that the base-change to $W(\mathsf{k})$ is again crucial in order to avoid the problem of conjugation by $\mathrm{Gal}(\mathsf{k}/\mathbf{F}_p)$. Thus, we have:

Proposition III.4.8 *For any residual parabolic eigenform $\bar{f} \in \mathbf{V}(\mathsf{k},\mathrm{N})$, there exists a complete local $W(\mathsf{k})$-algebra $\mathbf{R}(\bar{f})$ and a map*

$$\mathbf{f} : \mathbf{T}_0^\star(W(\mathsf{k}),\mathrm{N}) \longrightarrow \mathrm{R}(\bar{f})$$

[3] In the classical case, any algebra homomorphism from the restricted Hecke algebra to a p-adic ring arises by restriction (not only from a $\mathbf{Z}_p$-module homomorphism, but) from an *algebra* homomorphism from the full Hecke algebra to B (as we remark below, this follows from Atkin-Lehner theory). To put it another way, given a classical modular form which is an eigenform for the action of $\mathbf{T}_0^\star$, one can find an *eigenform* for all of $\mathbf{T}_0$ which "has the same q-expansion outside $\mathrm{N}p$". It is interesting to ask if this is still true in the p-adic case. The author does not know the answer except for the case $\mathrm{N} = 1$, in which the construction in Section II.3.3 shows that there exists an algebra homomorphism $\mathbf{T}_0 \longrightarrow \mathbf{T}_0^\star$ mapping U to 0, so that the answer is yes.

which is universal for the deformations outside $\mathrm{N}p$ *of* $\bar{f}$; *that is, so that for every* $f \in \mathbf{V}(B, \mathrm{N})$ *which is an eigenform outside* $\mathrm{N}p$ *and whose reduction modulo the maximal ideal is equal to* $\bar{f}$ *outside* $\mathrm{N}p$, *there exists a continuous homomorphism* $\alpha : \mathrm{R}(\bar{f}) \longrightarrow B$ *such that the map* $\phi_f^\star : \mathbf{T}_0^\star \longrightarrow B$ *defined by* f *is obtained by* $\phi_f^\star = \alpha \circ \mathbf{f}$.

The ring $\mathbf{R} = \mathbf{R}(\bar{f})$ will have an important role in what follows, especially in relation to constructing modular deformations of a residual representation, and we will need to recall one way to construct it.

For this, recall that $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$ is given as an inverse limit of Hecke algebras of finite rank; for any $\nu \geq 1$,

$$\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}) = \varprojlim_j \mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p^\nu),$$

where as above $\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p^\nu)$ denotes the restricted Hecke algebra corresponding to the space $\mathrm{S}^j(W(\mathsf{k}), \mathrm{N}p^\nu)$ of divided congruences of cuspforms of level $\mathrm{N}p^\nu$ and weight less than or equal to j defined over $W(\mathsf{k})$. Since the topology on $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$ is the inverse limit topology, and k is discrete, the map $\overline{\varphi}^\star$ necessarily factors through some $\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p^\nu)$, and hence also through any $\mathrm{h}_i^\star(W(\mathsf{k}), \mathrm{N}p^\nu)$ for any $i \geq j$; let $m(j)$ denote the kernel of the homomorphism $\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p^\nu) \longrightarrow \mathsf{k}$ corresponding to $\overline{\varphi}^\star$, for each j for which it exists. Since these j form an indexing set that is cofinal with the original one, we may as well take the inverse limit only over such j, without changing anything. We clearly have $\mathbf{m}^\star = \varprojlim_j m(j)$ and that, after localizing, a $W(\mathsf{k})$-algebra homomorphism

$$R \longrightarrow \mathbf{R} = \varprojlim_j \left(h_j^\star(\mathbf{Z}_p, \mathrm{N}p^\nu)\right)_{m(j)},$$

where R denotes the localization of $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$ at the maximal ideal $\mathbf{m}^\star$ and

$$\left(\mathrm{h}_j^\star(\mathbf{Z}_p, \mathrm{N}p^\nu)\right)_{m(j)}$$

denotes the localization of $\mathrm{h}_j^\star(\mathbf{Z}_p, \mathrm{N}p^\nu)$ at the maximal ideal $m(j)$.

Lemma III.4.9 *The inverse limit*

$$\mathbf{R} = \varprojlim_j \left(h_j^\star(\mathbf{Z}_p, \mathrm{N}p^\nu)\right)_{m(j)}$$

is the completion of $R = (\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}))_{\mathbf{m}^\star}$ *with respect to the topology induced by the inverse limit topology on* $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$.

Proof: This is easy to see, since the kernels of the maps

$$(\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}))_{\mathbf{m}^\star} \longrightarrow \left(\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p^\nu)\right)_{m(j)}$$

are clearly a basis of neighborhoods of zero in the induced topology. See also [EGA, 0_{I}.7.6]. $\square$

To summarize what has been accomplished in this section, given a residual eigenform $\bar{f}$, consider the corresponding algebra homomorphism $\bar{\varphi} : \mathbf{T}_0(W(\mathsf{k}), \mathrm{N}) \longrightarrow \mathsf{k}$ and its restriction $\bar{\varphi}^\star$ to $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$; then we may *identify* the three sets

$$\{\text{eigenforms outside } \mathrm{N}p \text{ defined over } B \text{ deforming } \bar{f} \text{ outside } \mathrm{N}p\}/_{\sim},$$

where we set $f \sim g$ if they are equal outside $\mathrm{N}p$,

$$\{W(\mathsf{k})\text{-algebra homomorphisms } \varphi^\star : \mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}) \longrightarrow B \text{ lifting } \bar{\varphi}^\star\},$$

and

$$\mathrm{Hom}_{alg}(\mathbf{R}(\bar{f}), B).$$

This allows us to think of the formal spectrum of $\mathbf{R} = \mathbf{R}(\bar{f})$ as the space of eigenforms outside $\mathrm{N}p$ deforming $\bar{f}$. The point of the next section is to identify this last set with the set of modular representations into $\mathrm{GL}_2(B)$ deforming the representation associated to $\bar{f}$. We will do this directly by constructing a representation associated to the "universal" object of the second set given by $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}) \longrightarrow \mathbf{R}$, which of course corresponds to the identity map of $\mathbf{R}$ in the third set. Before we do so, we need to recall a little of the classical theory of the restricted versus the unrestricted Hecke algebra.

III.4.3 Some classical results

In the classical situation, the Atkin-Lehner theory of "newforms" tells us that one can find a basis of $S(\overline{\mathbf{Q}}_p, j, \mathrm{N}p)$ (where the bar denotes algebraic closure) composed of "newforms" f_i and of "packages of oldforms"

$$\{V_d(g_i)\},$$

where g_i is a newform of level N_1 dividing $\mathrm{N}p$, d runs over the divisors of $\mathrm{N}p/\mathrm{N}_1$, and V_d denotes the map

$$S(\overline{\mathbf{Q}}_p, j, \mathrm{N}p/d) \longrightarrow S(\overline{\mathbf{Q}}_p, j, \mathrm{N}p)$$

given on q-expansions by $q \mapsto q^d$. (If $d \neq p$, V_d gives a map $\mathbf{V}(B, \mathrm{N}/d) \longrightarrow \mathbf{V}(B, \mathrm{N})$; of course, if $d = p$, V_d is just Frob.) The newforms f_i are eigenforms for $\mathbf{T}_0^\star$ "with multiplicity one", and are hence eigenforms for all of $\mathbf{T}_0$. As to the rest, every form in each "package" is an eigenform for $\mathbf{T}_0^\star$, with the *same* eigenvalues (i.e., $\mathbf{T}_0^\star$ acts as scalars). It is *not* always possible to diagonalize the action of $\mathbf{T}_0$ on each package of oldforms (though it is hard to come by an example!), but there is always at least one form in the space generated by the package which *is* an eigenform for $\mathbf{T}_0$ (simply because every matrix has at least one eigenvalue).

Let us call the basis we have been describing the Atkin-Lehner basis of $S(\overline{\mathbf{Q}}_p, k, \mathrm{N}p)$; assume such a basis has been chosen for each weight k. Taking a direct sum, we get a basis of

$$\mathsf{S}^k(\overline{\mathbf{Q}}_p, \mathrm{N}p) = \bigoplus_{i \leq k} S(\overline{\mathbf{Q}}_p, i, \mathrm{N}p).$$

Then, if we denote the Hecke algebra of $\mathrm{S}^k(\mathbf{Q}_p, \mathrm{N}p)$ by $\mathfrak{h}_k(\mathbf{Q}_p, \mathrm{N}p)$ and the corresponding restricted Hecke algebra by $\mathfrak{h}_k^\star(\mathbf{Q}_p, \mathrm{N}p)$, the choice of an Atkin-Lehner basis for each weight defines inclusions

$$\mathfrak{h}_k(\mathbf{Q}_p, \mathrm{N}p) \hookrightarrow \bigoplus K_i \oplus \bigoplus M_{n_j}(K_j)$$

and

$$\mathfrak{h}_k^\star(\mathbf{Q}_p, \mathrm{N}p) \hookrightarrow \bigoplus K_i \oplus \bigoplus K_j,$$

where, in each sum, i runs over the newforms, j runs over the packages of oldforms, n_j is the number of forms in the j^{th} package of oldforms, and K_i and K_j are finite extensions of $\mathbf{Q}_p$. The image of $\mathfrak{h}_k(\mathbf{Q}_p, \mathrm{N}p)$ is a commutative subalgebra of $\bigoplus K_i \oplus \bigoplus M_{n_j}(K_j)$, which may or may not be contained (up to conjugation) in $\bigoplus K_i \oplus \bigoplus (K_j)^{n_j}$, depending on whether the action of $\mathbf{T}_0$ may be diagonalized. It is interesting to consider the images of $\mathfrak{h}_k(\mathbf{Z}_p, \mathrm{N}p)$ and of $\mathfrak{h}_k^\star(\mathbf{Z}_p, \mathrm{N}p)$ under these inclusions. In fact, there is very little that is known beyond the (obvious) fact that these are orders in the images of the Hecke algebras over $\mathbf{Q}_p$, so that, for example, the image of $\mathfrak{h}_k^\star(\mathbf{Z}_p, \mathrm{N}p)$ is contained in a product of discrete valuation rings. This is the fact we will use later. One should note that, though we have stated everything for Hecke algebras over $\mathbf{Z}_p$ and $\mathbf{Q}_p$, the shape of the theory is the same for Hecke algebras over $W(\mathsf{k})$ and its field of fractions (just tensor).

III.5 Deformations of Galois Representations

We now go on to consider Galois representations and their deformations. As in the preceding chapter, we let k denote a finite field, and let $\mathcal{C}$ and $\hat{\mathcal{C}}$ be the categories defined above. Let $\mathcal{G}$ denote the absolute Galois group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ of the field of rational numbers.

Let B be a p-adic ring. We consider Galois representations "defined over B", i.e., representations

$$\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(B).$$

We will always assume ρ to be semisimple and unramified outside $\mathrm{N}p$. For any $\ell \nmid \mathrm{N}p$, let Φ_ℓ denote a Frobenius element for ℓ in $\mathcal{G}$; since ρ is unramified outside $\mathrm{N}p$, $\mathrm{trace}(\rho(\Phi_\ell))$ and $\det(\rho(\Phi_\ell))$ are well defined. We will say a given Galois representation is *modular* of tame level N if there exists a Katz eigenform $f \in \mathbf{V}(B, \mathrm{N})$ defined over B with q-expansion $f(q) = \sum a_i q^i$ (recall that by definition $a_1 = 1$) such that:

$$\mathrm{trace}(\rho(\Phi_\ell)) = a_\ell \quad \text{and} \quad \det(\rho(\Phi_\ell)) = \frac{1}{\ell}\chi(\langle \ell, \ell \rangle),$$

where of course χ denotes the character of $G(\mathrm{N})$ corresponding to f (its "weight and nebentypus" character). We say that the representation ρ is attached to the eigenform f (though it would be more precise to say it is attached to the system of eigenvalues outside $\mathrm{N}p$ corresponding to f). We will say a representation defined over $\mathbf{R}$ is a *family of modular representations* if it is attached, in the same sense as above, to a family of eigenforms $\mathbf{f}$. In this case it is clear that every specialization $\mathbf{R} \longrightarrow A$ will define a modular representation attached to the corresponding specialization of $\mathbf{f}$.

A *residual modular representation* will be a modular representation defined over a finite field k. Since any residual Katz eigenform f must be an eigenform for the diamond operators, we have $f \in \mathbf{V}_{1,1}$, and hence we know that f is a "modular form mod p" in the sense of Serre and Swinnerton-Dyer, i.e., that it is the reduction of a classical modular form of level N and some weight j. (Since residual eigenforms lift to classical eigenforms (see ahead), we know that a residual modular representation will be unramified outside Np.)

It is reasonable to look for modular representations because of the following well-known theorem (whose proof unfortunately is not, as far as this writer is aware, available in complete form anywhere in the literature):

Theorem III.5.1 *Let f be a classical modular form of level* $\mathrm{N}p^\nu$ *and weight $j \geq 2$ which is an eigenform under the action of the restricted Hecke algebra. Let χ denote the character describing the action of the diamond operators, and, for each $\ell \nmid \mathrm{N}p$, let a_ℓ be the eigenvalue corresponding to the Hecke operator T_ℓ. Let K be a finite extension of $\mathbf{Q}$ containing the a_ℓ and the values of χ. Let $\mathcal{P}$ be a place of K, of residual characteristic p, and let $K_{\mathcal{P}}$ be the completion of K at $\mathcal{P}$. Then there exists a continuous linear semi-simple representation*

$$\rho_f : \mathcal{G} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(K_{\mathcal{P}}),$$

which is unramified outside Np *and satisfies*

$$\mathrm{trace}(\rho_f(\Phi_\ell)) = a_\ell \quad \textit{and} \quad \det(\rho_f(\Phi_\ell)) = \frac{1}{\ell}\chi(\langle \ell, \ell \rangle),$$

for all $\ell \nmid \mathrm{N}p$, where as above Φ_ℓ denotes a Frobenius element for ℓ.

Proof: This is Theorem 6.1 of [DS74], where one may also find remarks on the extent to which proofs have been published. One should note that, since $\mathcal{G}$ is compact, it is always possible to modify ρ_f so that the image of $\mathcal{G}$ is contained in the ring of integers of $K_{\mathcal{P}}$. □

This shows that, at least for classical eigenforms f defined over valuation rings, one can find a representation ρ attached to f. To get the same result for classical *residual* eigenforms $\overline{f}$, one uses a lifting lemma:

Lemma III.5.2 *Let $\mathcal{H}$ be an arbitrary commutative algebra, R be a discrete valuation ring, and write $R\mathcal{H}$ for $R \otimes_{\mathbf{Z}} \mathcal{H}$. Let A and B be $R\mathcal{H}$-modules, and let $f : A \longrightarrow B$ be a surjective $R\mathcal{H}$-module homomorphism. Let $\Phi : \mathcal{H} \longrightarrow R$ be any map, and assume that there exists $v \in B$ such that $Tv = \Phi(T)v$ for any $T \in \mathcal{H}$. Let Q be a prime ideal in the support of Rv. Then there exists a discrete valuation ring R' of finite type over R and a map $\Psi : \mathcal{H} \longrightarrow R'$ such that:*

i. *there exists* $w \in A \otimes_R R'$ *such that* $Tw = \Psi(T)w$ *for any* $T \in \mathcal{H}$, *and*

ii. $\Psi(T) \equiv \Phi(T) \pmod{\mathcal{Q}'}$ *for all* $T \in \mathcal{H}$, *where* $\mathcal{Q}'$ *is the unique prime ideal of* R' *for which* $\mathcal{Q}' \cap R = \mathcal{Q}$.

Proof: This is [AS86, Prop. 1.2.2], and it is a generalization of [DS74, Lemma 6.11], which treats the case where A is a free R-module of finite rank and B is its reduction modulo the maximal ideal. □

To apply this to our situation, take $A = M(R, k, \mathrm{N})$, $B = M(R, k, \mathrm{N}) \otimes \mathbf{k} = M(R, k, \mathrm{N})$ (provided $k \neq 1$), and let $\mathcal{H}$ be the Hecke algebra corresponding to A. We get:

Corollary III.5.3 *Let* R *be a discrete valuation ring, and let* $\mathfrak{m} \subset R$ *be its maximal ideal. Suppose a classical modular form* $f \in M(R, k, \mathrm{N})$ *is an "normalized eigenform modulo* $\mathfrak{m}$*", i. e., that* $a_1(f) \equiv 1 \pmod{\mathfrak{m}}$ *and that, for any* $\mathrm{T} \in \mathcal{H}$, *we have*

$$\mathrm{T}f \equiv \lambda_{\mathrm{T}} f \pmod{\mathfrak{m}}.$$

Then there exists a classical modular form $g \in M(R, k, \mathrm{N})$ *which is an eigenform for the Hecke algebra and which satisfies* $g \equiv f \pmod{\mathfrak{m}}$.

Proof: The lemma says that one may find an eigenform g satisfying $\mathrm{T}g = \mu_{\mathrm{T}} g$ with $\mu_{\mathrm{T}} \equiv \lambda_{\mathrm{T}} \pmod{\mathfrak{m}}$. By the classical theory, we may assume g to be normalized, in which case a congruence of eigenvalues implies a congruence of q-expansions, and we are done. □

Note that the fact that the modular forms in question are *classical* is crucial to this result, since we need to work with R-modules of finite rank.

Now, using the preceding theorem, we get:

Proposition III.5.4 *Given a residual (normalized) eigenform* $f \in \mathbf{V}(\mathrm{k}, \mathrm{N})$, *there exists a residual Galois representation* ρ *unramified outside* $\mathrm{N}p$ *attached to it.*

Proof: Since any residual eigenform is necessarily classical of level N, this follows immediately from Lemma III.5.2 and Theorem III.5.1. (See [DS74, Theorem 6.7].) □

We will frequently write $\bar{f}$ for the residual eigenform and $\bar{\rho}$ for the residual representation under consideration, to emphasize that these objects are defined over $\mathbf{k}$. In what follows, A will always denote an object of the category $\hat{\mathcal{C}}$, hence in particular a noetherian local $W(\mathbf{k})$-algebra with residue field $\mathbf{k}$.

We now want to consider deformations of a residual modular representation $\bar{\rho}$ of (tame) level N and defined over $\mathbf{k}$, which we will take as fixed. Since every representation coming into consideration will then be unramified outside $\mathrm{N}p$, let us *redefine* $\mathcal{G}$ to be the Galois group of the maximal extension of $\mathbf{Q}$ unramified outside $\mathrm{N}p$ (i.e., the quotient of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ through which our representations factor).

We follow Mazur's definitions in [Ma], saying that two continuous homomorphisms $\mathcal{G} \longrightarrow \mathrm{GL}_2(A)$ are *strictly equivalent* if they differ by conjugation by an element of $\mathrm{GL}_2(A)$ which is in the kernel of the (residue) map $\mathrm{GL}_2(A) \longrightarrow \mathrm{GL}_2(\mathsf{k})$, and defining a *representation* to be a strict equivalence class of continuous homomorphisms, which we will nevertheless denote by $\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(A)$.

We will say a Galois representation $\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(A)$ is a *deformation to A* of a given residual representation $\overline{\rho}$ if any homomorphism $\mathcal{G} \longrightarrow \mathrm{GL}_2(A)$ in the strict equivalence class corresponding to ρ reduces to $\overline{\rho}$ under the canonical map $\mathrm{GL}_2(A) \longrightarrow \mathrm{GL}_2(\mathsf{k})$. (Clearly, this condition depends only on the strict equivalence class, and not on the specific homomorphism.) With these definitions Mazur has shown that *if $\overline{\rho}$ is absolutely irreducible*, there exists a complete noetherian local ring $\mathcal{R}$ with residue field k (so an object of $\hat{\mathcal{C}}$) and a representation $\rho \longrightarrow \mathrm{GL}_2(\mathcal{R})$ which is a universal deformation of $\overline{\rho}$, i.e., such that *any* deformation of $\overline{\rho}$ defined over a ring A in $\hat{\mathcal{C}}$ is obtained via a (unique) map $\mathcal{R} \longrightarrow A$. (For the construction in much greater generality, and many results about the ring $\mathcal{R}$, see [Ma].)

The point of this section is to construct a universal family of modular deformations of an absolutely irreducible residual representation $\overline{\rho}$, i.e., a representation

$$\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(\mathrm{R}),$$

where R is some complete noetherian local topological $\mathbf{Z}_p$-algebra in $\hat{\mathcal{C}}$, such that *any* modular deformation of $\overline{\rho}$ defined over a ring A in $\hat{\mathcal{C}}$ is obtained via a (unique) map $\mathrm{R} \longrightarrow A$. We will show that we may take R equal to the ring

$$\mathbf{R} = \mathbf{R}(\overline{f}) = \varprojlim_{i} (\mathrm{h}_i^\star(W(\mathsf{k}), \mathrm{N}p))_{m(i)}$$

considered in the previous section, which is the completion of the localization of the restricted Hecke algebra $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N})$ at the maximal ideal corresponding to the residual eigenform attached to $\overline{\rho}$.

We will construct our representation by using the classical representations, Proposition III.3.2, and a result on representations due to Mazur (in [Ma]). Let $\overline{\rho}$ be a residual modular representation, and let $\overline{f}$ be the associated residual eigenform. *In everything that follows, we assume the residual representation $\overline{\rho}$ to be absolutely irreducible.*

We will now construct a representation of $\mathcal{G}$ defined over the inverse limit $\mathbf{R}$, by constructing a representation defined over each of the localized restricted Hecke algebras at finite levels. As we saw above, the classical theory of newforms implies that we have an inclusion

$$(\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p))_{m(j)} \hookrightarrow \prod \mathcal{O}_i,$$

where the $\mathcal{O}_i$ are complete discrete valuation rings (with residue field k). (Localizing simply chooses some of the valuation rings obtained above, so that the residue fields are all the same and the residue maps all induce $\overline{\varphi}^\star$.) Composing with the projections, we get $W(\mathsf{k})$-algebra homomorphisms

$$(\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p))_{m(j)} \longrightarrow \mathcal{O}_i,$$

which by (III.1.3) and Atkin-Lehner theory (as above) correspond to normalized classical eigenforms of weight less than or equal to j (since a divided congruence which is an eigenform for the diamond operators must be a classical modular form of some weight), and hence to Galois representations $\mathcal{G} \longrightarrow \mathrm{GL}_2(\mathcal{O}_i)$. Taking the product, we get a Galois representation

$$\rho(j) : \mathcal{G} \longrightarrow \mathrm{GL}_2(\prod \mathcal{O}_i),$$

which satisfies, for $\ell \nmid \mathrm{N}p$,

$$\mathrm{trace}(\rho(j)(\Phi_\ell)) = t_\ell \quad \text{and} \quad \det(\rho(j)(\Phi_\ell)) = \frac{1}{\ell}\langle \ell, \ell \rangle,$$

where t_ℓ denotes the image of T_ℓ in $(h_j^\star(W(\mathsf{k}), \mathrm{N}p))_{m(j)}$, so that the traces of the representation we have constructed actually take values in the subalgebra $(\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p))_{m(j)}$ of $\prod \mathcal{O}_i$. It is at this point that we use Mazur's result.

Theorem III.5.5 *Let $A' \subset A$ be an inclusion of complete semi-local noetherian $W(\mathsf{k})$-algebras, with A' local. Let $A = \prod A_i$ be the factorization of A into a product of local rings, and assume that A' and all the A_i have residue field k. Let $\mathsf{r} : \mathcal{G} \longrightarrow \mathrm{GL}_2(A)$ be a continuous homomorphism such that the induced residual representations $\bar{\mathsf{r}}_i : \mathcal{G} \longrightarrow \mathrm{GL}_2(A_i) \longrightarrow \mathrm{GL}_2(\mathsf{k})$ are all equivalent and absolutely irreducible. Let H be the image of one of the $\bar{\mathsf{r}}_i$, and suppose that $\mathrm{H}^1(\mathrm{H}, \mathrm{Ad}^0_{\mathrm{H}}) = 0$, where $\mathrm{Ad}^0_{\mathrm{H}}$ denotes the k-vector space $\mathrm{M}_2(\mathsf{k})^0$ of two-by-two matrices of trace zero over k, endowed with the adjoint action of H.*

Suppose the traces of $\mathsf{r}(g)$ lie in A' for all $g \in \mathcal{G}$. Then there exists a continuous homomorphism $\mathsf{r}' : \mathcal{G} \longrightarrow \mathrm{GL}_2(A')$ such that the representation in $\mathrm{GL}_2(A)$ induced by the inclusion $A' \subset A$ is A-equivalent to r.

Proof: This is [Ma, §9 ,Cor. 1']. □

Theorem III.5.6 *Suppose that $p \geq 7$. Let $\bar{\rho}$ be an absolutely irreducible modular residual representation defined over k and of tame level N, and fix notations as above. Let $\mathbf{R}$ denote the inverse limit of the localizations of the restricted Hecke algebras at finite levels at the maximal ideals corresponding to the residual eigenform associated to $\bar{\rho}$:*

$$\mathbf{R} = \varprojlim_n (\mathrm{h}_j^\star(W(\mathsf{k}), \mathrm{N}p))_{m(j)}.$$

Then there exists a deformation

$$\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(\mathbf{R})$$

of $\bar{\rho}$ to $\mathbf{R}$, satisfying

$$\mathrm{trace}(\rho(\Phi_\ell)) = \mathrm{T}_\ell \quad \textit{and} \quad \det(\rho(\Phi_\ell)) = \frac{1}{\ell}\langle \ell, \ell \rangle$$

for any ℓ not dividing $\mathrm{N}p$, i.e., which is a family of modular representations of tame level N attached to the family of eigenforms (outside $\mathrm{N}p$) defined by the canonical homomorphism $\mathbf{T}_0^\star \longrightarrow \mathbf{R}$.

Proof: Given what we have already done, it is only necessary to note that the hypothesis on $\mathrm{H}^1(\mathrm{H}, \mathrm{Ad}^0_{\mathrm{H}})$ is satisfied when $p \neq 5$ (see [CPS]). This gives representations into $\mathrm{GL}_2((h_j(W(\mathsf{k}), \mathrm{N}p))_{m(j)})$ for all sufficiently large j. Taking the inverse limit then gives ρ; the properties of traces and determinants follow at once from those of the $\rho(j)$, and we are done. □

Corollary III.5.7 *Under the hypotheses of the Theorem, the ring* $\mathbf{R} = \mathbf{R}(\bar{f})$ *is noetherian.*

Proof: It is clearly a quotient of Mazur's ring $\mathcal{R}$ which gives the universal (modular or not) deformation of $\bar{\rho}$, and which is noetherian by construction. □

Corollary III.5.8 *Let $\bar{f}$ be a residual eigenform whose associated Galois representation is absolutely irreducible, let A be a p-adic ring with residue field k, and let $f \in \mathbf{V}_{par}(B, \mathrm{N})$ be any eigenform outside $\mathrm{N}p$ over A which reduces to $\bar{f}$ outside $\mathrm{N}p$. Let χ be the weight-and-nebentypus character of f (i.e., the character giving the action of $G(\mathrm{N})$). Then there exists a representation*

$$\rho_f : \mathcal{G} \longrightarrow \mathrm{GL}_2(A)$$

attached to f, i.e., satisfying, for each $\ell \nmid \mathrm{N}p$,

$$\mathrm{trace}(\rho_f(\Phi_\ell)) = a_\ell \quad \textit{and} \quad \det(\rho_f(\Phi_\ell)) = \frac{1}{\ell}\chi(\langle \ell, \ell \rangle).$$

Proof: The eigenform f corresponds to a continuous homomorphism $\mathbf{T}_0^\star \longrightarrow A$, which determines a continuous homomorphism $\mathbf{R} \longrightarrow A$ mapping T_ℓ to a_ℓ and $\langle \ell, \ell \rangle$ to $\chi(\langle \ell, \ell \rangle)$. Composing this map with ρ gives the desired representation. (Note that this result was previously only known for *classical* eigenforms.) □

Conversely, if a Galois representation deforming an absolutely irreducible residual representation $\bar{\rho}$ to a p-adic ring A is attached to an eigenform $f \in \mathbf{V}_{par}(B, \mathrm{N})$, then it is (up to strict equivalence) perforce obtained from ρ via the map $\mathbf{R} \longrightarrow A$ induced by the map $\mathbf{T}_0^\star(W(\mathsf{k}), \mathrm{N}) \longrightarrow A$ corresponding to the eigenform. Thus, ρ is the universal family of (level N) modular deformations of $\bar{\rho}$. We call the ring $\mathbf{R}$ the universal (level N) modular deformation ring of $\bar{\rho}$; when the dependence on $\bar{\rho}$ or on $\bar{f}$ must be made explicit, we will write $\mathbf{R} = \mathbf{R}(\bar{\rho}) = \mathbf{R}(\bar{f})$.

Caveat: It is important to note that this construction is strongly dependent on the "tame level". It takes account of all deformations arising from modular forms of level $\mathrm{N}p^\nu$ for any ν, but it does *not* cover those arising, say, from modular forms of level $\mathrm{N}^2 p$ (which are still unramified outside $\mathrm{N}p$!). The whole question of the effect of changing the level on our construction is a difficult one, and we will remark further on questions of this type in the conclusion of this chapter.

If we denote Mazur's universal deformation ring by $\mathcal{R} = \mathcal{R}(\overline{\rho})$, it follows that we have a tautological epimorphism

$$\mathcal{R}(\overline{\rho}) \longrightarrow \mathbf{R}(\overline{\rho}).$$

One suspects that this is in fact an isomorphism, i.e., that every deformation of a modular residual representation is modular. We want, in the next section, to obtain information about the modular deformation space and its relation to the total deformation space.

III.6 The modular deformation space

In [Ma], Mazur has examined in some detail the structure of the space of deformations of a residual representation. His approach is to view a representation defined over a p-adic ring B (deforming a given residual representation) as a B-valued point in the formal scheme $\mathrm{Spf}(\mathcal{R}(\overline{\rho}))$. The problem then gets translated into studying the geometry of that formal scheme. The surjection $\mathcal{R}(\overline{\rho}) \longrightarrow \mathbf{R}(\overline{\rho})$ should then be viewed as defining the closed formal subscheme $\mathrm{Spf}(\mathbf{R}(\overline{\rho}))$ of modular deformations of the given residual representations, and one would like to obtain more information about this space and its inclusion in the space of all deformations.

III.6.1 Changing the weight

The first thing to note is that, even staying within the theory of classical modular forms, there are many deformations of any given residual representation. Let $\overline{f}$ be a residual eigenform, and let $\overline{\rho}$ be the attached representation. Assume that $\overline{f}$ is p-adically of weight k, with $0 \leq k < p-1$, i.e., that it transforms under $(\mathbf{Z}/p\mathbf{Z})^\times$ via the k^{th} power of the Teichmüller character. Then we know (Lemma III.5.2 and the fact that residual eigenforms are necessarily classical) that there exists a classical eigenform f_i of level N and weight $i \equiv k \pmod{p-1}$ lifting $\overline{f}$; choose one such with minimal i.

For each $m \geq 0$, we may consider $\overline{f}$ as the reduction of the modular form $f_i \mathrm{E}_{p-1}^m$, which is of weight $j = i + m(p-1)$. Hence, by the lifting lemma (III.5.3), $\overline{f}$ can be lifted to an eigenform of weight j, level N, and defined over some extension of B, for each such j. Thus, we have shown:

Proposition III.6.1 *Suppose $\overline{f}$ is a residual eigenform which is the reduction of a classical modular $g_i \in M(B, i, \mathrm{N})$ (not necessarily an eigenform). Then, for each j satisfying $j \equiv i \pmod{p-1}$ and $j \geq i$, there exists a finite extension B_j of B and a classical eigenform $f_j \in M(B_j, j, \mathrm{N})$ which reduces to $\overline{f}$ outside* Np.

Of course, to each such f_j one can associate a representation ρ_j, which will be a deformation of the representation $\overline{\rho}$ attached to $\overline{f}$. This should be viewed as saying that one can vary the deformation "in the direction of the weight".

In fact, one can improve this result a little if one is willing to change the level. Recall, first, that there exists an Eisenstein series $\mathrm{E} \in M(B, 1, \mathrm{N}p)$ (i.e., of weight *one* on $\Gamma_1(\mathrm{N}p)$) satisfying

$$\mathrm{E}(q) \equiv 1 \pmod{p}.$$

Hence, if f_i is a lifting to weight i and level $\mathrm{N}p$, $\nu \geq 1$ we may multiply by E to get a form of weight $i+1$ and level $\mathrm{N}p$ which still reduces to $\bar{f}$. Applying Corollary III.5.3 once again gives:

Proposition III.6.2 *Let $\bar{f}$ be a residual eigenform, and assume that there exists a classical modular form $g_i \in M(B, i, \mathrm{N}p)$ reducing to $\bar{f}$. Then, for every $j \geq i$, there exists a finite extension B_j of B and an eigenform $f_j \in M(B_j, j, \mathrm{N}p)$ which reduces to $\bar{f}$.*

The p-part of the nebentypus of f_j is determined by the reduction $\bar{f}$, so that we know the precise p-adic weight of any deformation to classical forms of level $\mathrm{N}p$: if $\bar{f}$ is of p-adic weight k, i.e., if we have

$$\langle x, 1\rangle \bar{f} = \omega^k(x)\bar{f}$$

(in $\mathbf{V} \otimes \mathsf{k}$), then f_j must have p-adic weight (k, j), i.e.,

$$\langle x, 1\rangle f_j = \omega^k(x)\left(\frac{x}{\omega(x)}\right)^j = \omega^{k-j}(x)x^j.$$

In classical terms, the p-part of the nebentypus of a lifting of weight j must be ω^{k-j}.

Remark: It is of course possible to state a similar proposition for level $\mathrm{N}p^\nu$, but it is not clear that anything new would be gained. In other words, the forms of which we would be asserting the existence could simply be the $f_j \in M(B', j, \mathrm{N}p)$ (thought of as of level $\mathrm{N}p^\nu$ via the canonical inclusion). To get things which are indeed of higher level, we will need another method, which is the theme of the next subsection.

III.6.2 Twisting

In the classical theory of modular forms, one encounters the operation of "twisting by a Dirichlet character". If χ is a character modulo M taking values in a discrete valuation ring B and f is a modular form of level N with q-expansion $f(q) = \sum a_n q^n$, this produces a modular form f^χ of level NM^2 whose q-expansion is

$$f^\chi(q) = \sum \chi(n) a_n q^n,$$

where of course we extend χ to all of $\mathbf{Z}$ in the usual way. (We will recall the definition of f^χ below.) We will be interested in the case when M is a power of p, in which case χ may simply be thought of as a character of finite order of $\mathbf{Z}_p^\times$, i.e., $\chi : \mathbf{Z}_p^\times \longrightarrow B$ factoring through a quotient $(\mathbf{Z}/p^\nu\mathbf{Z})^\times$ (which we extend to all of $\mathbf{Z}_p$ by setting it equal to zero on $p\mathbf{Z}_p$).

Let $\mathbf{Z}_p^\times = (\mathbf{Z}/p\mathbf{Z})^\times \times \Gamma$ be the usual decomposition of $\mathbf{Z}_p^\times$. We will say that the character χ is wild if it is trivial on $(\mathbf{Z}/p\mathbf{Z})^\times$, i.e., if it is of order a power of p. In that case, its values will we p-power roots of unity in B, and hence will be congruent to 1 modulo the maximal ideal $\mathfrak{m}$ of B, so that we will have $f^\chi(q) \equiv f(q) \pmod{\mathfrak{m}}$. Since twisting by a character transforms eigenforms into eigenforms, this will give us a new source of deformations of a residual eigenform.

The point of this section is to recall the definition in modular terms of the operation of twisting by a character, and show that it can in fact be extended to any character of $\mathbf{Z}_p^\times$, of finite or infinite order. Let us first consider the classical case of a character of finite order. By the q-expansion principle, we may extend our ring as necessary, so that we may as well take B to be the completion of the ring of integers in a separable closure of $\mathbf{Q}_p$. A trivial check on q-expansions then allows us to determine over which ring our forms turn out to be defined.

Let χ be a character of $\mathbf{Z}_p^\times$ factoring through $(\mathbf{Z}/p^n\mathbf{Z})^\times$, and let $f \in \mathbf{V}(B,\mathrm{N})$. To define the twist $f^\chi \in \mathbf{V}(B,\mathrm{N})$, we must specify its value on a trivialized elliptic curve with an arithmetic level N structure. Let E/B be an elliptic curve, $\varphi : \hat{\mathrm{E}} \xrightarrow{\sim} \hat{\mathbf{G}}_m$ be a trivialization, and $\imath : \boldsymbol{\mu}_\mathrm{N} \hookrightarrow \mathrm{E}$ be an arithmetic level N structure. Then φ^{-1} determines (by restriction) an inclusion $\boldsymbol{\mu}_{p^n} \hookrightarrow \mathrm{E}$, and we consider the quotient $\mathrm{E}_1 = \mathrm{E}/\varphi^{-1}(\boldsymbol{\mu}_{p^n})$, with its induced trivialization φ_1 (if $\pi : \mathrm{E} \longrightarrow \mathrm{E}_1$ is the projection, $\varphi_1 = \varphi \circ \check{\pi}$, which is an isomorphism of formal groups because $\check{\pi}$ is étale). The image of $\mathrm{E}[p^n]$ (the kernel of p^n in E) in E_1 is canonically isomorphic to the constant group scheme $\mathbf{Z}/p^n\mathbf{Z}$ (via the Weil pairing), so that we have an isomorphism

$$\mathrm{E}_1[p^n] \cong \boldsymbol{\mu}_{p^n} \times \mathbf{Z}/p^n\mathbf{Z},$$

i.e., what Katz calls an arithmetic $\Gamma(p^n)$ structure on E_1 (see [Ka76, 2.3] for more details).

Suppose now that $\mathrm{H} \subset \mathrm{E}_1$ is an étale subgroup of order p^n. Then we may use the isomorphism above to associate to H a p^n-th root of unity ζ_H (in such a way that we associate 1 to the subgroup $\mathbf{Z}/p^n\mathbf{Z}$). Then we define:

Definition III.6.3 *Let $f \in \mathbf{V}(B,\mathrm{N})$, and let χ be a character of $\mathbf{Z}_p^\times$ factoring through $(\mathbf{Z}/p^n\mathbf{Z})^\times$. Then we define the twist of f by χ by*

$$f^\chi(\mathrm{E},\varphi,\imath) = \frac{1}{p^n} \sum_{x \bmod p^n} \sum_{\substack{\mathrm{H}\subset\mathrm{E}_1 \\ \#\mathrm{H}=p^n \\ \mathrm{H}\text{ étale}}} \chi(x)\zeta_\mathrm{H}^{-x} f(\mathrm{E}_1/\mathrm{H},\varphi_\mathrm{H},\imath_\mathrm{H}),$$

where E_1 is as above, and where φ_H and $\imath_\mathrm{H}$ are induced from φ and $\imath$ in the obvious way.

This defines f^χ as an element of $\mathbf{V}[\frac{1}{p}]$; to show that it is in fact in $\mathbf{V}$, we need only show that its q-expansion has integral coefficients. So we must evaluate f^χ on the Tate curve. Let $\mathrm{E} = \mathrm{Tate}(q)$, and let φ and $\imath$ be the canonical trivialization and level N structure. Then we have $\mathrm{E}_1 = \mathrm{Tate}(q^p)$, and the isomorphism obtained above is:

$$\begin{array}{ccc} \boldsymbol{\mu}_{p^n} \times \mathbf{Z}/p^n\mathbf{Z} & \longrightarrow & \mathrm{E}_1[p^n] \\ (\zeta, j) & \mapsto & \zeta q^j, \end{array}$$

where of course we think of $\mathrm{Tate}(q^p)$ as the quotient of $\mathbf{G}_m$ by $q^{p\mathbf{Z}}$. The étale subgroups of order p^n are then $\mathrm{H}_i = <\zeta^i q>$ (the group generated by $\zeta^i q$), where ζ is a generator of $\boldsymbol{\mu}_{p^n}$, and we simply have $\zeta_{\mathrm{H}_i} = \zeta^i$. Then it is easy to see that $\mathrm{E}_1/\mathrm{H}_i \cong \mathrm{Tate}(q)$, via the map $q \mapsto \zeta^i q$, and this isomorphism is compatible with the canonical trivialization and level N structure. Hence, if $f(q) = \sum a_n q^n$, we have:

$$\begin{aligned} f^{\chi}(\mathrm{E},\varphi,\imath) &= \frac{1}{p^n} \sum_{x \bmod p^n} \sum_{\substack{\mathrm{H}\subset \mathrm{E}_1 \\ \#\mathrm{H}=p^n \\ \mathrm{H}\ \text{étale}}} \chi(x)\zeta_{\mathrm{H}}^{-x} f(\mathrm{E}_1/\mathrm{H}, \varphi_{\mathrm{H}}, \imath_{\mathrm{H}}) \\ &= \frac{1}{p^n} \sum_{x \bmod p^n} \sum_{i \bmod p^n} \chi(x)\zeta^{-ix} \sum_n \zeta^{in} a_n q^n \\ &= \frac{1}{p^n} \sum_n \sum_{x \bmod p^n} \chi(x) \sum_{i \bmod p^n} \zeta^{i(n-x)} a_n q^n \\ &= \sum_n \chi(n) a_n q^n. \end{aligned}$$

This shows:

Proposition III.6.4 *Let $f \in \mathbf{V}(B,\mathrm{N})$, and let $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ be a character of finite order, which we extend to $\mathbf{Z}_p$ by setting it equal to zero on $p\mathbf{Z}_p$. Assume f has q-expansion $f(q) = \sum a_n q^n$; then there exists a p-adic modular function $f^\chi \in \mathbf{V}(B,\mathrm{N})$ whose q-expansion is given by*

$$f^\chi(q) = \sum \chi(n) a_n q^n.$$

Thus, we have defined, for each character of finite order χ of $\mathbf{Z}_p^\times$ with values in B, a B-linear endomorphism

$$\begin{array}{ccc} \mathbf{V}(B,\mathrm{N}) & \longrightarrow & \mathbf{V}(B,\mathrm{N}) \\ f & \mapsto & f^\chi. \end{array}$$

It is immediately clear that we always have $\mathrm{U} f^\chi = 0$ by construction; we will show ahead that whenever f is an eigenform, so is f^χ. First, however, we would like to extend this construction to more general characters, i.e., we would like to show that there exists a twist f^χ for any character $\mathbf{Z}_p^\times \longrightarrow B$. We do this by noting that the definition of $f \mapsto f^\chi$ can be interpreted as the integral of the character χ with respect to a certain measure on $\mathbf{Z}_p$ taking values in the space $\mathrm{End}_{\mathbf{Z}_p}(\mathbf{V}(\mathbf{Z}_p,\mathrm{N}))$.

With notations as above, we define a measure μ on $\mathbf{Z}_p$ as follows: for each $a \in \mathbf{Z}_p$ and each $n \geq 0$, consider the endomorphism $\mu(a,n)$ of $\mathbf{V}$ defined by

$$(\mu(a,n)f)(\mathrm{E},\varphi,\imath) = \frac{1}{p^n} \sum_{\substack{\mathrm{H}\subset \mathrm{E}_1 \\ \#\mathrm{H}=p^n \\ \mathrm{H}\ \text{étale}}} \zeta_{\mathrm{H}}^{-a} f(\mathrm{E}_1/\mathrm{H}, \varphi_{\mathrm{H}}, \imath_{\mathrm{H}}).$$

Then $\mu(a,n)f$ is clearly an element of $\mathbf{V}[\frac{1}{p}]$; to check that it is in fact in $\mathbf{V}$, we need only compute the effect on q-expansions. This is completely analogous to the preceding calculation, and we get: if $f(q) = \sum a_n q^n$, then

$$(\mu(a,n)f)(q) = \sum_{n \equiv a \pmod{p^n}} a_n q^n.$$

This shows:

Lemma III.6.5 *Let $f \in \mathbf{V} = \mathbf{V}(\mathbf{Z}_p, \mathrm{N})$; then*

i. for any $a \in \mathbf{Z}_p$ and any $n \geq 0$, $\mu(a,n)f \in \mathbf{V}$, and

ii. we have

$$\mu(a,n)f = \sum_{\substack{b \bmod p^{n+1} \\ b \equiv a \pmod{p^n}}} \mu(b, n+1)f.$$

Proof: Given the above computation, the first statement follows immediately from the q-expansion principle. It is enough to check the second statement on q-expansions, in which case it is obvious. □

Thus, for each $f \in \mathbf{V}$, the assignment

$$a + p^n\mathbf{Z}_p \longmapsto \mu(a,n)f$$

defines a $\mathbf{V}$-valued measure on $\mathbf{Z}_p$; varying f we get that the assignment

$$a + p^n\mathbf{Z}_p \longmapsto \mu(a,n)$$

defines a $\mathrm{End}_{\mathbf{Z}_p}(\mathbf{V})$-valued measure, which we will denote simply by μ. Of course, for any p-adic ring B, we may think of μ as taking values in $\mathrm{End}_B(\mathbf{V}(B,\mathrm{N}))$. Then it is clear that, for any character $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ of finite order, extended to $\mathbf{Z}_p$ by setting it equal to zero on $p\mathbf{Z}_p$, we have

$$f^\chi = \left(\int_{\mathbf{Z}_p} \chi d\mu\right) f,$$

so that the operation of "twisting by χ" is the integral of χ with respect to our measure μ. This allows us to extend the idea of twisting by a character to any character χ of $\mathbf{Z}_p^\times$:

Definition III.6.6 *Let $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ be any homomorphism, and let $f \in \mathbf{V}(B,\mathrm{N})$. Then we define the "twist of f by χ", denoted f^χ, by*

$$f^\chi = \left(\int_{\mathbf{Z}_p} \chi d\mu\right) f,$$

where as before we extend χ to $\mathbf{Z}_p$ by $\chi(p\mathbf{Z}_p) = 0$.

Of course it is just as easy to twist by any function: given a continuous function $\alpha : \mathbf{Z}_p \longrightarrow B$ we write

$$f^{\alpha} = \left(\int_{\mathbf{Z}_p} \alpha d\mu\right) f,$$

and sometimes refer to this as "the twist of f by α".

We would now like to understand the effect of twisting by a character on Katz eigenforms (or on eigenforms outside $\mathrm{N}p$). As we have already remarked, it is clear (from the q-expansions, for example) that we will have $\mathrm{U}(f^{\chi}) = 0$ for *any* f. However, this is simply an artifact of our construction (to be precise, it is a consequence of extending χ by zero to $\mathbf{Z}_p$), and the effect of the Hecke and diamond operators requires analysis. We consider first the case of a character of finite order.

Lemma III.6.7 *Let $\chi : \mathbf{Z}_p^{\times} \longrightarrow B^{\times}$ be a character of finite order, factoring through $(\mathbf{Z}/p^n\mathbf{Z})^{\times}$. Then we have, for any $f \in \mathbf{V}(B, \mathrm{N})$:*

$$\langle 1, y\rangle(f^{\chi}) = (\langle 1, y\rangle f)^{\chi}, \tag{III.2}$$

for any $y \in (\mathbf{Z}/\mathrm{N}\mathbf{Z})^{\times}$,

$$\langle y, 1\rangle(f^{\chi}) = \chi(y)^2(\langle y, 1\rangle f)^{\chi}, \tag{III.3}$$

for any $y \in \mathbf{Z}_p^{\times}$,

$$\mathrm{U}(f^{\chi}) = 0, \tag{III.4}$$

and

$$\mathrm{T}_{\ell}(f^{\chi}) = \chi(\ell)(\mathrm{T}_{\ell} f)^{\chi}, \tag{III.5}$$

for any prime $\ell \neq p$.

Proof: Since the definition of twisting by χ doesn't involve the level N structure $\imath : \boldsymbol{\mu}_{\mathrm{N}} \hookrightarrow \mathrm{E}$ in any way, the first equation is obvious.

For the second equation, recall that the action of the diamond operator $\langle y, 1\rangle$ is given by

$$(\langle y, 1\rangle f)(\mathrm{E}, \varphi, \imath) = f(\mathrm{E}, y^{-1}\varphi, \imath),$$

where we use the canonical action of $\mathbf{Z}_p^{\times}$ on $\hat{\mathbf{G}}_m$. Let, as above, $\mathrm{E}_1 = \mathrm{E}/\varphi^{-1}(\boldsymbol{\mu}_p)$. Then we need to compare the assignment $\mathrm{H} \mapsto \zeta_{\mathrm{H}}$ induced by φ to that induced by $y^{-1}\varphi$. The twisting by y^{-1} affects this in two ways: the inclusion $\boldsymbol{\mu}_p \longrightarrow \mathrm{E}_1$ gets twisted by y^{-1}, and so does the Weil pairing on $\mathrm{E}[p]$, and hence the identification $\mathrm{E}[p]/\varphi^{-1}(\boldsymbol{\mu}_p) \cong \mathbf{Z}/p\mathbf{Z}$. The net effect is changing ζ_{H} to $\zeta_{\mathrm{H}}^{y^{-2}}$. To conclude, we simply calculate, using the multiplicativity of χ:

$$\begin{aligned}
\langle y, 1\rangle(f^{\chi})(\mathrm{E}, \varphi, \imath) &= f^{\chi}(\mathrm{E}, y^{-1}\varphi, \imath) \\
&= \frac{1}{p^n}\sum_{x}\sum_{\mathrm{H}} \chi(x)(\zeta_{\mathrm{H}}^{y^{-2}})^{-x} f(\mathrm{E}_1/\mathrm{H}, y^{-1}\varphi, \imath) \\
&= \chi(y)^2 \frac{1}{p^n}\sum_{x}\sum_{\mathrm{H}} \chi(y^{-2}x)\zeta_{\mathrm{H}}^{-y^{-2}x}(\langle y, 1\rangle f)(\mathrm{E}_1/\mathrm{H}, \varphi, \imath) \\
&= \chi(y)^2(\langle y, 1\rangle f)^{\chi}(\mathrm{E}, \varphi, \imath).
\end{aligned}$$

Hence, we get $\langle y,1\rangle(f^\chi) = \chi(y)^2(\langle y,1\rangle f)^\chi$, as desired. Putting the first two equations together, we get, in particular, that, for any $\ell \nmid \mathrm{N}p$,

$$\langle \ell,\ell\rangle(f^\chi) = \chi(\ell)^2(\langle \ell,\ell\rangle f)^\chi.$$

To get the equation for the action of T_ℓ, we may either work as above or compute directly on q-expansions; we use the latter. Given any modular function $g \in \mathbf{V}(B,\mathrm{N})$, we write $a_n(g)$ for the coefficient of q^n in the q-expansion of g. Then we know that, if $\ell \nmid \mathrm{N}$,

$$a_n(\mathrm{T}_\ell g) = a_{n\ell}(g) + \frac{1}{\ell}a_{n/\ell}(\langle \ell,\ell\rangle g),$$

where we make the convention that $a_{n/\ell}(g) = 0$ if ℓ does not divide n; if $\ell|\mathrm{N}$, the equation becomes

$$a_n(\mathrm{T}_\ell g) = a_{n\ell}(g).$$

Recall that, for χ of finite order,

$$a_n(f^\chi) = \chi(n)a_n(f).$$

The case when $\ell|\mathrm{N}$ then follows immediately. For the other case,

$$\begin{aligned}
a_n(\mathrm{T}_\ell f^\chi) &= a_{n\ell}(f^\chi) + \frac{1}{\ell}a_{n/\ell}(\langle \ell,\ell\rangle(f^\chi)) \\
&= a_{n\ell}(f^\chi) + \frac{1}{\ell}a_{n/\ell}(\chi(\ell)^2(\langle \ell,\ell\rangle f)^\chi) \\
&= \chi(n\ell)a_{n\ell}(f) + \frac{1}{\ell}\chi(\ell)^2 a_{n/\ell}((\langle \ell,\ell\rangle f)^\chi) \\
&= \chi(n)\chi(\ell)a_{n\ell}(f) + \frac{1}{\ell}\chi(\ell)^2\chi(n/\ell)a_{n/\ell}(\langle \ell,\ell\rangle f) \\
&= \chi(\ell)\chi(n)\left(a_{n\ell}(f) + \frac{1}{\ell}a_{n/\ell}(\langle \ell,\ell\rangle f)\right) \\
&= \chi(\ell)a_n((\mathrm{T}_\ell f)^\chi).
\end{aligned}$$

By the q-expansion principle, it follows that $\mathrm{T}_\ell(f^\chi) = \chi(\ell)(\mathrm{T}_\ell f)^\chi$, and we are done. □

If we examine the calculations above, it is clear that we have only used the fact that $\chi : \mathbf{Z}_p \longrightarrow B$ is a (locally constant) multiplicative function. Let $\alpha : \mathbf{Z}_p \longrightarrow B$ be any multiplicative function, so that $\alpha(xy) = \alpha(x)\alpha(y)$ for any $x, y \in \mathbf{Z}_p$. If we approximate α by locally constant functions α_n, these will satisfy

$$\alpha_n(xy) \equiv \alpha_n(x)\alpha_n(y) \pmod{p^{\nu(n)}}$$

for some $\nu(n)$, so that the calculations above all go through after changing some of the equalities to congruences. Taking the limit, we get:

Corollary III.6.8 *Let $\alpha : \mathbf{Z}_p \longrightarrow B$ be any continuous multiplicative function, and let $f \in \mathbf{V}(B,\mathrm{N})$ with q-expansion $f(q) = \sum a_n q^n$. Then:*

i. $(f^\alpha)(q) = \sum \alpha(n) a_n q^n$,

ii. $\langle 1, y\rangle(f^\alpha) = (\langle 1, y\rangle f)^\alpha$, *for any* $y \in (\mathbf{Z}/\mathrm{N}\mathbf{Z})^\times$,

iii. $\langle y, 1\rangle(f^\alpha) = \alpha(\ell)^2(\langle y, 1\rangle f)^\alpha$, *for any* $y \in \mathbf{Z}_p^\times$,

iv. $\mathrm{U}(f^\chi) = \alpha(p)(\mathrm{U}f)^\alpha$ *and*

v. $\mathrm{T}_\ell(f^\alpha) = \alpha(\ell)(\mathrm{T}_\ell f)^\alpha$, *for any prime* $\ell \neq p$.

In particular, we get:

Corollary III.6.9 *Let* $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ *be any character, extended to* $\mathbf{Z}_p$ *by* $\chi(p\mathbf{Z}_p) = 0$, *and let* $f \in \mathbf{V}(B, \mathrm{N})$ *be a Katz eigenform with weight-and-nebentypus character* $\varepsilon : G(\mathrm{N}) \longrightarrow B^\times$ *and with* $\mathrm{T}_\ell f = \lambda_\ell f$ *for any prime* $\ell \neq p$. *Then* f^χ *is a Katz eigenform with weight-and-nebentypus character* $\varepsilon\chi^2$, *satisfying* $\mathrm{U}f^\chi = 0$ *and* $\mathrm{T}_\ell(f^\chi) = \chi(\ell)\lambda_\ell f^\chi$.

For some special functions $\alpha : \mathbf{Z}_p \longrightarrow B$, the operation of twisting by α turns out to be a well-known operator on modular functions. For a first example, let $\chi : \mathbf{Z}_p^\times \longrightarrow \mathbf{Z}_p^\times$ be the trivial character: $\chi(x) = 1$ for all $x \in \mathbf{Z}_p^\times$. Let $f \in \mathbf{V}$ have q-expansion $f(q) = \sum a_n q^n$. Then

$$f^\chi(q) = \sum_{(n,p)=1} a_n q^n = (f - \mathrm{Frob}(\mathrm{U}f))(q),$$

so that twisting by the trivial character is the same as $1 - \mathrm{Frob} \circ \mathrm{U}$. For a more interesting example, let $\alpha : \mathbf{Z}_p \longrightarrow \mathbf{Z}_p$ be the identity function. Then we get

$$f^\alpha(q) = \sum n a_n q^n = q\frac{d}{dq} f(q),$$

so that twisting by the identity is the same as the "$q\frac{d}{dq}$" operator considered by Serre (in [Se73]) and Katz (in [Ka76]). Since α is multiplicative, it follows from the lemma that $q\frac{d}{dq}$ "is of p-adic weight 2", i.e., sends modular forms of weight (i, k) to forms of weight $(i + 2, k + 2)$, which is a result obtained by Katz in [Ka76].

III.6.3 Families of twists, and an estimate for the Krull dimension of the modular deformation ring

We can now use the twisting process defined above to produce one-parameter families of deformations of a residual eigenform (or, equivalently, of a modular residual deformation). If we begin with a Katz eigenform f with weight-and-nebentypus character $\varepsilon : \mathbf{Z}_p^\times \longrightarrow B^\times$, then twisting by a character $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ produces another Katz eigenform f^χ, with weight-and-nebentypus character $\epsilon\chi^2$. In particular, if χ is wild, i.e., if $\chi(x)$ is a one-unit in $B^\times$ for all $x \in \mathbf{Z}_p^\times$, then we will have $f^\chi \equiv f \pmod{\mathfrak{m}}$, where $\mathfrak{m}$ is the maximal ideal of B. Thus, if f is a deformation of a residual eigenform $\bar{f}$, we have constructed a family of deformations of $\bar{f}$ indexed by the wild characters $\mathbf{Z}_p^\times \longrightarrow B^\times$, where of course we may vary B. In what follows, we will always assume B to be a finite extension of $\mathbf{Z}_p$.

To actually construct a family of functions (as defined above) recall the identifications of $\mathbf{V}(B,\mathrm{N})$ with $\mathrm{Hom}_{\mathbf{Z}_p,cont}(\mathbf{T}_0(W(\mathsf{k}),\mathrm{N}),B)$ (Section III.1) and of $\Lambda_B = B[[\Gamma]]$ with the space of B-valued measures on Γ (see, e.g., [Se73]). Let $\mathbf{\Lambda}_B = B[[\mathbf{Z}_p^\times]] = \Lambda_B[(\mathbf{Z}/p\mathbf{Z})^\times]$, which we of course can identify with the space of B-valued measures on $\mathbf{Z}_p^\times$. Given a Katz eigenform $f \in \mathbf{V}(B,\mathrm{N})$, consider the $\mathbf{V}(B,\mathrm{N})$-valued measure on $\mathbf{Z}_p^\times$ given by (the restriction of) $\mu(f)$. For each $\mathrm{T} \in \mathbf{T}_0$, evaluation at T gives a B-valued measure on $\mathbf{Z}_p^\times$, hence an element of $\mathbf{\Lambda}_B$. This defines a continuous homomorphism $\tau_f : \mathbf{T}_0(W(\mathsf{k}),\mathrm{N}) \longrightarrow \mathbf{\Lambda}_B$, and hence a family of modular forms. In explicit terms, τ_f can be described as follows: if

$$\phi_f : \mathbf{T}_0 \longrightarrow B$$

is the canonical homomorphism corresponding to f, we have:

i. $\tau_f(\langle 1,y\rangle) = \phi_f(\langle 1,y\rangle)$,

ii. $\tau_f(\langle y,1\rangle) = \phi_f(\langle y,1\rangle)\langle y,1\rangle^2$,

iii. $\tau_f(\mathrm{U}) = 0$, and

iv. $\tau_f(\mathrm{T}_\ell) = \phi_f(\mathrm{T}_\ell)\langle \ell,1\rangle$.

Then, for any character χ, f^χ is just the specialization of the family τ_f via χ (and, in fact, the result that τ_f is a continuous homomorphism follows at once from this fact, which is obvious from the q-expansions). If f is a deformation of a residual eigenform $\bar{f}$, and if we restrict to wild characters, which of course are trivial on $(\mathbf{Z}/p\mathbf{Z})^\times \subset \mathbf{Z}_p^\times$, we get a continuous homomorphism $\mathbf{T}_0(W(\mathsf{k}),\mathrm{N}) \longrightarrow \Lambda_B$, i.e., a family of modular forms, all of which reduce to $\bar{f}$ outside $\mathrm{N}p$ (because $\chi(x) \equiv 1 \pmod{p}$ for all x. Thus, we get a continuous homomorphism $\tau_f^\star : \mathbf{T}_0^\star(W(\mathsf{k}),\mathrm{N}) \longrightarrow \Lambda_B$ which defines a continuous homomorphism $\mathbf{R} \longrightarrow \Lambda_B$ (because every f^χ reduces to $\bar{f}$). Since $\Lambda_B \cong B[[T]]$, this should be thought of as a one-dimensional analytic family of deformations of $\bar{f}$.

To summarize:

Proposition III.6.10 *Given any residual eigenform $\bar{f}$ of level* N*, and given any Katz eigenform $f \in \mathbf{V}(B,\mathrm{N})$ reducing to $\bar{f}$ modulo the maximal ideal of B, there exists a one-dimensional analytic family of deformations*

$$\tau_f^\star : \mathbf{R} \longrightarrow \Lambda_B$$

giving the twists of f by wild characters of $\mathbf{Z}_p^\times$.

In terms of representations, if $\bar{\rho}$ is the residual representation attached to $\bar{f}$, we get, by composing the universal modular deformation $\rho : \mathcal{G} \longrightarrow \mathrm{GL}_2(\mathbf{R})$ we get, for every modular deformation $\rho_f : \mathcal{G} \longrightarrow \mathrm{GL}_2(B)$, a one-dimensional analytic family of deformations $\mathcal{G} \longrightarrow \mathrm{GL}_2(\Lambda_B)$, whose specialization under any character $\chi : \Gamma \longrightarrow B^\times$ is simply the twist of ρ_f by the one dimensional representation of $\mathcal{G}$ given by χ in the obvious way. (See Mazur's paper [Ma] for more on twisting representations.) Thus we have shown that any twist of a modular deformation of $\bar{f}$ by a wild character (i.e., a character of Γ) is again a modular deformation of $\bar{f}$. We will use this to show that the Krull dimension of the universal modular deformation ring is at least 3. For this, we need two easily-proved facts:

Lemma III.6.11 *Let $f \in \mathbf{V}(B, \mathrm{N})$ be a deformation of a residual eigenform $\bar{f}$, and let $\tau_f^\star : \mathbf{R}_B = \mathbf{R}\hat{\otimes}B \longrightarrow \Lambda_B$ be the map induced by the family of twists of f defined above. Then $\tau_f^\star$ is a continuous surjective homomorphism of $\mathbf{Z}_p$-algebras.*

Proof: Everything but surjectivity has already been noted. To get surjectivity, it is enough to check that, for any $\gamma \in \Gamma$, the element $\langle\gamma\rangle$ is in the image (where, as above, we use angular brackets to distinguish elements of Γ from themselves as elements of $\mathbf{Z}_p^\times$). But, since 2 is invertible in $\mathbf{Z}_p$ and so is the image γ under the character corresponding to f, this is immediate from the formula $\tau_f^\star(\langle y, 1\rangle) = \phi_f(\langle y, 1\rangle)\langle y, 1\rangle^2$ above. In fact, this shows that the composite map $\Lambda_B \longrightarrow \mathbf{R}_B \longrightarrow \Lambda_B$ is already surjective. □

Lemma III.6.12 *Let $f \in \mathbf{V}(B, \mathrm{N})$ be a deformation of a residual eigenform $\bar{f}$, and assume f is classical, i.e., $f \in M(B, k, \epsilon, \mathrm{N}p^\nu)$ for some ν and appropriate nebentypus ϵ. Then*

i. the twist of f by any character of finite order is again classical; specifically, if $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ factors through $(\mathbf{Z}/p^n\mathbf{Z})^\times$, we have $f^\chi \in M(B, k, \epsilon\chi^2, \mathrm{N}p^{\nu+2n})$

ii. the twist of f by any character $\chi : \mathbf{Z}_p^\times \longrightarrow B^\times$ of infinite order is not *a classical modular form.*

Proof: The first claim, is, of course, well-known, and obvious from the definitions. For the second claim, we must use considerations of p-adic Hodge structure of the corresponding representations. One knows that, in the classical case, the p-adic Hodge twists (in the sense of [MW86]) of a representation attached to a classical form of weight k whose reduction is absolutely irreducible are $(0, k-1)$ (see [MW86]); twisting by the character $\gamma \mapsto \gamma^j$ gives a representation with twists $(j, j+k-1)$, which is therefore not classical. □

Then we get:

Proposition III.6.13 *The Krull dimension of the universal modular deformation space $\mathbf{R} = \mathbf{R}(\bar{f}) = \mathbf{R}(\bar{\rho})$ is at least three.*

Proof: Let k be the least weight for which there is a classical lift f_k of level N. Then for each $j \geq k$, $j \equiv k \pmod{p-1}$, there exists a classical lift f_j of weight j and level N. Then, for each such j, we have obtained a continuous surjective homomorphism

$\tau_f^\star : \mathbf{R} \longrightarrow \Lambda_{B_j}$ (where we B_j is a finite extension of $\mathbf{Z}_p$). Since the dimension of Λ_{B_j} is two, we need only prove that the kernel of one (and hence of any) of these maps is not a minimal prime ideal of $\mathbf{R}$. However, the preceding corollary shows that these ideals are all distinct; since $\mathbf{R}$ is noetherian, they cannot be minimal primes (because there is only a finite number of such), and we are done. □

It would be interesting to get a more precise estimate of the Krull dimension of the modular deformation ring. For example, Mazur and Boston have considered the case of a residual representation

$$\overline{\rho} : \mathcal{G} \longrightarrow S_3 \subset \mathrm{GL}_2(\mathbf{F}_p),$$

where S_3 is the symmetric group on three letters. Under some additional hypotheses, they show that the universal deformation space $\mathcal{R}(\overline{\rho})$ has Krull dimension four (in fact, that it is a power-series ring in three variables over $\mathbf{Z}_p$). Is the modular deformation ring also of dimension four in this case?

III.6.4 The ordinary case

In the case where the residual representation is attached to an ordinary modular form, one can consider the "universal ordinary deformation", i.e., the universal deformation associated to ordinary modular forms. Using the work of Hida in [Hi86a], Mazur and Wiles (in [MW86]) constructed the universal ordinary deformation (under some restrictive hypotheses) and determined several of its properties, especially with respect to the image of the decomposition group at p. In this section, we summarize these results and point out their relation to our larger deformation space.

For simplicity, and to agree with the situation in [MW86], let us assume that the level $\mathrm{N} = 1$. Let $\bar{f} \in \mathbf{V}_{par}(\mathsf{k}, 1)$ be an ordinary residual parabolic eigenform, and let $\overline{\rho}$ be the attached residual representation. By duality, $\bar{f}$ corresponds to a map $\mathbf{T}_0 \longrightarrow \mathsf{k}$, which necessarily factors through the ordinary Hecke algebra $\mathbf{T}_0^{ord} = e_0\mathbf{T}_0$. In fact, we can say more; a residual eigenform has a weight $\chi : \mathbf{Z}_p^\times \longrightarrow \mathsf{k}$, which, as we remarked above, must be a power of the Teichmüller character: $\chi = \omega^i$. Then it is clear that the map $\phi_{\bar{f}} : \mathbf{T}_0 \longrightarrow \mathsf{k}$ corresponding to $\bar{f}$ must in fact factor through the summand T_i of $\mathbf{T}_0^{ord}$ defined in the Appendix to the last chapter. The various Hecke algebras and the map defined by $\bar{f}$ fit together like this:

$$\mathbf{T}_0^\star \subset \mathbf{T}_0 \xrightarrow{e_0} \mathbf{T}_0^{ord} \longrightarrow \mathsf{T}_i \longrightarrow \mathsf{k}.$$

Let $\mathbf{m}^\star$, $\mathbf{m}$, $\mathbf{m}^{ord}$, and $\mathfrak{m}$ denote the kernels of $\phi_{\bar{f}}$ in each of the Hecke algebras.

The construction of Mazur and Wiles requires a hypothesis on the weight i. Hence *assume, for the remainder of this section, that* $i \neq 2$. In terms of the representation $\overline{\rho}$, this means that $det(\overline{\rho})$ is not the cyclotomic character.

In this situation, Mazur and Wiles construct a representation

$$\rho^{ord} : \mathcal{G} \longrightarrow \mathrm{GL}_2(\mathsf{R}),$$

where R denotes the completion of the Hecke algebra T_i at the maximal ideal m, which gives the universal ordinary deformation of the representation attached to $\bar{f}$, in the same sense as before: any representation deforming $\bar{\rho}$ which is attached to an ordinary p-adic modular form is obtained from ρ^{ord} via the induced map from R to the ring of definition of the attached modular form. They then obtain theorems about the action the decomposition and the inertia groups at p, including a description of the p-adic Hodge twists of the specializations. We refer to [MW86] for further details. One should note that Mazur and Wiles have also obtained a necessary condition, in terms of the action of the inertia group at p, representation to be attached to an ordinary modular form. This condition is conjectured to be sufficient, in which case it would give representation-theoretic description of the ordinary deformation space. See the discussion in [MW86] and [Ma] for more details.

We would like to compare the universal ordinary deformation to the universal modular deformation we constructed before. It is clear, from the universal property of the constructions, that there must exist a map $\mathbf{R} \longrightarrow \mathsf{R}$ so that ρ^{ord} is the representation obtained from $\boldsymbol{\rho}$ via this map; it is also clear (compare the traces of the Frobenii) that this map must be the map induced by the canonical maps

$$\mathbf{T}_0^\star \hookrightarrow \mathbf{T}_0 \longrightarrow \mathsf{T}_i$$

by completion at the maximal ideal. From the analysis in [MW86], we can show:

Proposition III.6.14 *Let $\bar{f}$ be an ordinary residual eigenform of p-adic weight $i \neq 1$ and level $\mathrm{N} = 1$. Let $\mathbf{R} = \mathbf{R}(\bar{f})$ be the universal modular deformation ring and let $\mathsf{R} = \mathsf{R}(\bar{f}) = (\mathsf{T}_i)_{\mathsf{m}}$ be the completion of the i^{th} component of the ordinary Hecke algebra at the ideal corresponding to $\bar{f}$. Consider the map $\mathbf{R} \longrightarrow \mathsf{R}$ induced from the composition of the inclusion with the projection on T_i. Then R is at most a quadratic extension of the image of $\mathbf{R}$.*

Proof: Since we are assuming $\mathrm{N} = 1$, it is clear that R is generated over the image of $\mathbf{R}$ by (the image in the completion of) the U operator. To see that U is quadratic over the image of $\mathbf{R}$, we show that for some unit $\lambda \in \Lambda$, the element $\mathrm{U} + \lambda \mathrm{U}^{-1} \in \mathsf{R}$ belongs to the image of $\mathbf{R}$. (Note that U is invertible in R, by definition of the ordinary part.) This, however, follows at once from the results in [MW86, §8] mentioned above: let σ be an element of the decomposition group at p mapping modulo the inertia group to a generator of $\hat{\mathbf{Z}} = \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$; then we have

$$\mathrm{trace}(\rho^{ord}(\sigma)) = \mathrm{U} + \lambda \mathrm{U}^{-1},$$

where λ is some unit of Λ; therefore, for any such σ,

$$\mathrm{trace}(\boldsymbol{\rho}(\sigma)) \mapsto \mathrm{U} + \lambda \mathrm{U}^{-1} \in \mathsf{R},$$

so that $\mathrm{U} + \lambda \mathrm{U}^{-1}$ is in the image of $\mathbf{R}$, as desired. □

Since R is a flat Λ-algebra of finite rank (see Section II.4), its Krull dimension is two. Taking all possible twists gives a map

$$\tau^{ord} : \mathbf{R} \longrightarrow \mathsf{R}[[\Lambda]],$$

as in the preceding section (because a twist of an ordinary modular form is never ordinary). This last ring has Krull dimension three, so that we need to study the kernel $\mathcal{I} = \ker(\tau^{ord})$; if this is not contained in a minimal prime ideal of $\mathbf{R}$, it will follow that the dimension of $\mathbf{R}$ is at least four. In the example worked out by Mazur and Boston ("neat S_3 extensions"), this would suffice to show that every deformation is modular.

We have not been able to prove that the dimension of the modular deformation space is at least four. In any case, it is easy to see that the modular deformation space is always strictly larger than the ordinary deformation space, by showing that there exist deformations of the representation $\overline{\rho}$ attached $\overline{f}$ which are modular, but whose corresponding modular forms are neither ordinary nor twists of ordinary modular forms. Assume $k \geq 3$, $k \equiv i \pmod{p-1}$. Then $\overline{f}$ can be lifted to a classical modular form of weight k which can be assumed to be an eigenform and is necessarily ordinary (because it is congruent to $\overline{f}$). For deforming the *representation*, however, all we need is to find a lift that is congruent to $\overline{f}$ *outside* $\mathrm{N}p$; we will show that if the weight is high enough, one can always find a deformation outside $\mathrm{N}p$ which is not ordinary, but is classical of level N, so that it is certainly not a twist of a classical modular form.

Let f be any lift of $\overline{f}$ to an (ordinary) eigenform of weight k and level N. Assume for simplicity that $\mathsf{k} = \mathbf{F}_p$ and therefore that f is defined over some totally ramified extension B of $\mathbf{Z}_p$. Consider the modular form

$$g_1 = \mathrm{E}_{p-1}^{j}(\mathrm{E}_{p-1}^{k} f - (\mathrm{T}_p f)^p) \in M(B, pk + j(p-1), \mathrm{N}).$$

It is clear that the reduction of g_1 is an eigenform (since it is congruent, modulo the maximal ideal of B, to $f - \mathrm{Frob}(\mathrm{U}f)$); thus, by Lemma III.5.2, there must exist an eigenform $g \in M(B, pk + j(p-1), \mathrm{N})$ (possibly after base-change) which is congruent to g_1. Then g will be a deformation of $\overline{f}$ outside $\mathrm{N}p$, it will be classical of weight $pk + j(p-1)$ and level N, and it will not be ordinary. Hence the representation attached to g will not be obtained by specializing the universal ordinary representation ρ^{ord}.

Thus, we have shown:

Proposition III.6.15 *Let $\overline{f}$ be an ordinary residual eigenform defined over* k. *Suppose there exists a classical eigenform f of weight k_0 and level* N *defined over a $W(\mathsf{k})$-algebra B reducing to $\overline{f}$ modulo the maximal ideal, and hence ordinary:* $\mathrm{T}_p f \not\equiv 0 \pmod{\mathfrak{m}}$. *Then, for any weight $k \geq pk_0$, $k \equiv k_0 \pmod{p-1}$, there exists a classical eigenform g of weight k and level* N *defined over a finite extension B' of B (which will depend on k) whose reduction modulo the maximal ideal is equal to $\overline{f}$ outside* $\mathrm{N}p$ *but which is* not *ordinary:* $\mathrm{T}_p g \equiv 0 \pmod{\mathfrak{m}}$.

Remark: To be absolutely precise, f itself is not ordinary, because it is not an eigenform under the U operator. However, its ordinary projection $e_0 f$ is nonzero and equal to f outside p, while $e_0 g = 0$. The point is simply that for each large enough weight, there is always at least one ordinary and one non-ordinary deformation which is not the twist of any ordinary deformation.

Corollary III.6.16 *If the residual representation $\overline{\rho}$ is attached to an ordinary residual eigenform $\overline{f}$, the modular deformation space of $\overline{\rho}$ is strictly larger than the ordinary deformation space, even if we add all the twists of ordinary deformations.*

Since the Krull dimension of the ordinary-plus-twists deformation ring is three, one is led to ask:

Question III.1 *Suppose that the residual modular representation $\overline{\rho}$ is attached to an* ordinary *modular form. Is the Krull dimension of the universal modular deformation ring always greater than or equal to four?*

At least in the case when $\overline{f}$ is ordinary, the preceding discussion suggests that the answer may be yes; in the "neat S_3" case considered by Mazur and Boston, this would imply that every deformation of the residual representation in question is attached to a p-adic modular form.

III.7 Further Questions

This final section collects some of the questions which seem to arise in relation to the topics we have discussed in this book, and that may suggest paths for further research. When appropriate, we have suggested what we suspect will be the answer, but we have not dignified these suspicions by calling them conjectures, since for the most part there is little evidence one way or the other. We use the notations and conventions defined above, giving references only when necessary.

To begin with, there are several questions associated to the subject of Chapter II: the U operator and its eigenforms. As we saw, there exists, for each $\alpha \geq 0$ a "slope α projection"

$$e_\alpha : \mathsf{M}^\dagger(B,k,\mathrm{N};1)\otimes K \longrightarrow \mathsf{M}^\dagger(B,k,\mathrm{N};1)\otimes K,$$

defining a splitting

$$\mathsf{M}^\dagger(B,k,\mathrm{N};1)\otimes K = M^{(\alpha)} \oplus F^{(\alpha)},$$

where $M^{(\alpha)}$ is the "slope α eigenspace", i.e., the finite-dimensional subspace spanned by the generalized eigenforms for the U operators corresponding to eigenvalues with valuation α. As we have seen, e_α is a continuous linear endomorphism of the p-adic Banach space $\mathsf{M}(B,k,\mathrm{N};r)\otimes K$ (which is given the p-adic topology for which the B-module $\mathsf{M}(B,k,\mathrm{N};r)$ is the closed unit ball), but it is *not* continuous on $\mathsf{M}^\dagger(B,k,\mathrm{N};1)\otimes$

K when this is given the q-expansion topology (except when $\alpha = 0$). The first important question, then, which we have already mentioned above, is whether for fixed r the norm of e_α is bounded independent of the weight k.

Question III.2 *Let $e_\alpha^{(k)}$ denote the slope α projection on $\mathsf{M}(B,k,\mathrm{N};r)\otimes K$. Does there exist a bound $C(\alpha,r)$ (independent of k) so that $\|e_\alpha\| \leq C(\alpha,r)$? If so, can we take $C(\alpha) = p^\alpha$?*

As we remarked in Chapter II, the answer to both questions is yes when $\alpha = 0$; we conjectured above (Conjecture II.4) that the bound $C(\alpha,r)$ always exists, but we have no idea whether it is plausible that $C(\alpha,r) = p^\alpha$ in general.

Closely related are the following two questions:

Question III.3 *Suppose $f \in \mathsf{M}^\dagger(B,k,\mathrm{N};1)$ satisfies $e_\alpha f = f$; so that in particular we have $f \in \mathsf{M}(B,k,\mathrm{N};r)\otimes K$ for any r such that* $\mathrm{ord}(r) < p/(p+1)$. *If k is large enough, do we in fact have $f \in M(B,k,\mathrm{N}) \otimes K = M(K,k,\mathrm{N})$? In other words, is it true that any overconvergent modular form of slope α and sufficiently high weight is necessarily a* classical *modular form?*

If $\alpha = 0$, the answer is yes, and "sufficiently high" is $k \geq 3$, as was shown by Hida (see Section II.4 above). If f is an *eigenform*, we already know that f must be *congruent* to a classical modular form; if we can show that this classical modular form must also be of slope α, this would answer the question affirmatively in the case of eigenforms.

Finally, we have:

Question III.4 *Let $P_k(t)$ denote the characteristic power series of the U operator acting on $\mathsf{M}^\dagger(B,k,\mathrm{N};1)\otimes K$. Is it true that if $k_1 \equiv k_2 \pmod{p^{n-1}(p-1)}$ then $P_{k_1}(t) \equiv P_{k_2}(t) \pmod{p^n}$? If so, is the variation in fact locally analytic in k?*

Above, we conjectured that the answer to the first part of this question is "yes", but made a guess that the answer to the second question is "no", unless we consider truncations of the full characteristic power series at the various slopes.

Going on to questions suggested by the work in this chapter, the most natural and important question has already been stated:

Question III.5 *Let $\overline{\rho}$ be a residual Galois representation attached to a residual eigenform $\overline{f}$. Is it true that every deformation of $\overline{f}$ is attached to a p-adic modular form reducing to $\overline{f}$?*

Less ambitiously, one could ask

Question III.6 *Let $\overline{\rho}$ be a residual Galois representation attached to a residual eigenform $\overline{f}$. What is the Krull dimension of the modular deformation ring $\mathbf{R}(\overline{f})$?*

We have shown that this Krull dimension is at least three, and we have given an argument that suggests that when $\bar{f}$ is *ordinary* (i.e., of slope 0) this dimension should in fact be at least four. It is not clear to what extent this dimension will depend on $\bar{f}$ (rather than, say, only on its weight and level).

Another category of question relating to the construction of the Galois representations attached to the various deformations of a residual eigenform has to do with the relation between representation-theoretic properties of the representation deforming $\bar{\rho}$ and the properties of the modular form attached to it. For example, Mazur and Wiles have shown that any representation attached to an ordinary modular form must satisfy a certain condition on the action of an inertia group at p (see [MW86]).

Question III.7 *Let $\bar{\rho}$ be an absolutely irreducible residual representation attached to a residual eigenform $\bar{f}$. Let ρ be any modular deformation, and let f be the p-adic modular form attached to it. Can one find a condition on ρ that will hold if and only if f is overconvergent? Is it true that f is ordinary if and only if ρ satisfies the condition of Mazur and Wiles?*

There are several other crucial questions about the modular deformation space which have been touched upon only lightly in this chapter; they have to do with the dependence of our construction on the level. As we pointed out above, we have always assumed that the level N of our modular forms was fixed beforehand, and that all modular forms under consideration were (p-adically) of level N. (Recall that this includes classical modular forms of level $\mathrm{N}p^{\nu}$ for every ν; in general, we may always assume that all levels are prime to p, since introducing powers of p does not alter the spaces of p-adic modular forms in question.) In fact, there are two possible ways to vary the level: adding new prime divisors to the level or not.

To begin with, let N_1 be a number prime to p, and let N be the product of its prime divisors (this is sometimes called the *radical* of N_1). It is clear that we have an inclusion $\mathbf{V}(B,\mathrm{N}) \hookrightarrow \mathbf{V}(B,\mathrm{N}_1)$, and hence that we have an epimorphism of Hecke algebras $\mathbf{T}(B,\mathrm{N}_1) \longrightarrow \mathbf{T}(B,\mathrm{N})$; localizing and completing at the maximal ideal corresponding to some eigenform $\bar{f} \in \mathbf{V}(\mathsf{k},\mathrm{N})$, we get a map between the modular deformation ring of level N_1 and the modular deformation ring of level N:

$$\mathbf{R}^{(\mathrm{N}_1)}(\bar{f}) \longrightarrow \mathbf{R}^{(\mathrm{N})}(\bar{f}).$$

Note that these are both deformation rings for the representation associated to $\bar{f}$, since "unramified outside $\mathrm{N}p$" and "unramified outside $\mathrm{N}_1 p$" are synonymous.

Question III.8 *What is the relation between these two rings of deformations outside $\mathrm{N}p$? Do they have the same Krull dimension? Can they be distinguished representation-theoretically, say, by a finer examination of the ramification at the primes dividing* N*?*

We may, of course, take the inverse limit over all N_1 with the same prime divisors, and get a deformation of the representation associated to $\bar{f}$ to an even larger ring

$$\mathbf{R}^{(\infty)}(\bar{f}) = \varprojlim_{\mathrm{N}_1} \mathbf{R}^{(\mathrm{N}_1)}(\bar{f}).$$

Question III.9 *What are the properties of this larger deformation ring* $\mathbf{R}^{(\infty)}(\bar{f})$*? In particular, what is its Krull dimension? Is it equal to Mazur's full deformation ring? Is it equal to the modular deformation ring of level* N_1 *if* N_1 *is sufficiently large?*

In the case where N|M and there are primes dividing M which do not divide M, the situation is even more interesting. If we consider a residual eigenform $\bar{f} \in \mathbf{V}(B,\mathrm{N})$, we may think of it as being of level N, in which case we will look for representations unramified outside Np, *or* of level M, in which case we will look for representations unramified outside Mp, which form a larger space. Hence we get a map of deformation spaces

$$\mathbf{R}^{(\mathrm{M})}(\bar{f}) \longrightarrow \mathbf{R}^{(\mathrm{N})}(\bar{f}),$$

which we may think of as defining the space of deformations unramified outside Np as a subscheme of the space of deformations unramified outside Mp.

Question III.10 *What is the relation between the spaces* $\mathbf{R}^{(\mathrm{M})}(\bar{f})$ *and* $\mathbf{R}^{(\mathrm{N})}(\bar{f})$*? Do they have the same Krull dimension?*

Even more interesting is the following situation: suppose $\overline{\rho}$ is a residual Galois representation unramified outside Np and which is known to be attached to a residual eigenform of level Mp, for some M as above. Then we can consider three deformation spaces: first, Mazur's deformation ring $\mathcal{R}^{(\mathrm{M})}(\overline{\rho})$ of $\overline{\rho}$ *considered as unramified outside* Mp (so we look for deformations deforming $\overline{\rho}$ which are unramified outside Mp), second, the ring $\mathbf{R}^{(\mathrm{M})}(\overline{\rho})$, corresponding to (level M) *modular* deformations, and third, the ring $\mathcal{R}^{(\mathrm{N})}(\overline{\rho})$, corresponding to deformations (modular or not) which are unramified outside Np. We have a diagram of surjections:

$$\begin{tikzcd}
 & \mathcal{R}^{(\mathrm{N})}(\overline{\rho}) \\
\mathcal{R}^{(\mathrm{M})}(\overline{\rho}) \arrow[ur] \arrow[dr] & \\
 & \mathbf{R}^{(\mathrm{M})}(\overline{\rho})
\end{tikzcd}$$

We can interpret this as defining two subspaces of the space of deformations unramified outside Mp: the subspace of level M modular deformations and the subspace of deformations unramified outside Np.

Question III.11 *What is the intersection of these two subspaces? In particular, is it always non-empty? In other words, does the existence of deformations unramified outside* Np *imply the existence of* modular *deformations which are of level* N*?*

Making N precise is of course crucial here. If this is done correctly, an affirmative answer is of course expected, since the problem can be reworded as a part of a well-known conjecture due to Serre:

Conjecture III.1 *Suppose $\overline{\rho}$ is a residual representation attached to a residual eigenform $\bar{f}$ of (weight k and) squarefree level* M*, and suppose that $\overline{\rho}$ is in fact unramified at some prime ℓ dividing* M*. Then there exists a residual eigenform of (possibly different weight and) level* $\mathrm{N} = \mathrm{M}/\ell$ *to which $\overline{\rho}$ is attached. Put in other words, if $\bar{f}$ is the reduction of a classical modular form of level* M *whose attached representation is unramified outside* M/ℓ*, then there exists a classical modular form of level* M/ℓ *(but possibly of different weight) which also reduces to $\bar{f}$.*

Serre's conjecture, of course, goes on to define precisely the minimal possible weight for the lifting. It seems to us that the "level part" of the conjecture should be accessible by p-adic methods (the "weight part" is probably not). One might attempt, for example, to prove it first in the ordinary case.