$\mathbf{R} = \mathbf{T}$ for Artin characters

Grant T. Barkley

In previous talks we have discussed the statement of modularity lifting and introduced the $\mathbf{R} = \mathbf{T}$ formalism. The modularity theorem and Taylor–Wiles modularity lifting are both statements about proving modularity of 2-dimensional representations. We discussed how modular forms correspond to automorphic representations of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, and how the modularity theorem is a special case of the Langlands correspondence for GL_2, which in this case asserts that every Galois representation $\mathrm{Gal}(\overline{\mathbb{Q}} \mid \mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Q}_\ell)$ satisfying certain properties should have an associated modular form. Today we will demonstrate the main points of the Taylor–Wiles proof in a simpler setting: we will prove the modularity theorem for GL_1, or, equivalently, the Kronecker–Weber theorem.

Disclaimer: I am not an expert, and there are likely to be mistakes throughout this note. I have tried to point out some subtleties that I found while learning the material presented here, but their multitude suggests that I have likely missed some. Reader beware! My primary references for the GL_2 side of things are [FLT] and [DDT] from the seminar webpage. Those should be taken as the more definitive source.

1 Between GL_1 and GL_2: a dictionary

1.1 Shimura varieties

1.1.1. Let us recall some setup and context for the GL_2 theory we have discussed so far. We defined the (open) modular variety $Y(N)$ to a variety over $\mathbb{Q}$ whose complex points are a quotient space for the action of a congruence subgroup $\Gamma(N)$ on the upper half-plane $\mathcal{H}$. The (compactified) modular curve $X(N)$ is formed by adding a finite set of cusps to $Y(N)$. The curve $Y(N)$ is a **Shimura variety** for GL_2. This means that its complex points can be realized adelically as the double coset space

$$Y(N)(\mathbb{C}) = \mathrm{GL}_2(\mathbb{Q}) \backslash \mathcal{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/\widehat{\Gamma}(N).$$

Here $\mathbb{A}_f$ denotes the finite adeles of $\mathbb{Q}$, which are isomorphic to $\widehat{\mathbb{Z}} \otimes \mathbb{Q}$.

1.1.2. For GL_1 we have an analogous story. The open modular variety $Y(N)$ should have complex points which are a nice quotient, and since we will see that $Y(N)$ is already proper over $\mathbb{Q}$, the compactified modular variety $X(N)$ is the same as $Y(N)$. We can also view the complex points of $Y(N)$ as an adelic double coset space:

$$Y(N)(\mathbb{C}) = \mathrm{GL}_1(\mathbb{Q}) \backslash \{\pm 1\} \times \mathrm{GL}_1(\mathbb{A}_f)/\widehat{\Gamma}(N).$$

Here $\mathbb{Q}^\times$ acts on $\{\pm 1\}$ via the sign representation and $\widehat{\Gamma}(N)$ is the kernel of $\mathrm{GL}_1(\widehat{\mathbb{Z}}) \twoheadrightarrow \mathrm{GL}_1(\mathbb{Z}/N)$. We can actually compute the right-hand side explicitly: each $\mathbb{Q}^\times$ coset of $\{\pm 1\} \times \mathrm{GL}_1(\mathbb{A}_f)$ contains a unique element of $\mathrm{GL}_1(\widehat{\mathbb{Z}})$, so the full double quotient is just $\mathrm{GL}_1(\mathbb{Z}/N) = (\mathbb{Z}/N)^\times$. This tells us $Y(N)$ is some variety over $\mathbb{Q}$ whose complex points are in bijection with $(\mathbb{Z}/N)^\times$. We also want $Y(N)$ to be connected and 0-dimensional. There is a natural candidate: $\operatorname{Spec} \mathbb{Q}(\zeta_N)$. The complex points of this variety are equivalently field embeddings $\mathbb{Q}(\zeta_N) \hookrightarrow \mathbb{C}$. The group $\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})$ acts simply transitively on these embeddings, so if we pick an embedding $\mathbb{Q}(\zeta_N) \hookrightarrow \mathbb{C}$ then we get an identification

$$(\operatorname{Spec} \mathbb{Q}(\zeta_N))(\mathbb{C}) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q}) = (\mathbb{Z}/N)^\times.$$

Hence we will take $X(N) = Y(N) = \operatorname{Spec} \mathbb{Q}(\zeta_N)$ as our modular variety. This is also the canonical example of a **0-dimensional Shimura variety**.

1.2 Coherent cohomology

1.2.1. For GL_2 we have been interested in the vector space of modular forms of a given weight and level. We saw that weight k forms of level N live in a tensor power of the Hodge bundle on the universal elliptic curve over $X(N)$. For the modularity of elliptic curves, we are especially interested in the space $S_2(N;\mathbb{C})$ of weight 2 cusp forms of level N. (Really those of level $\Gamma_0(N)$, or equivalently those with the correct conductor and nebuntypus, but these differences will be invisible for GL_1.) For a subring R of $\mathbb{C}$, we let $S_2(N;R)$ denote the subset of $S_2(N;\mathbb{C})$ consisting of modular forms whose q-expansion has its coefficients living in R. Then

$$S_2(N;\mathbb{Q}) = \mathsf{H}^0(X(N), \Omega_{X(N)/\mathbb{Q}}) = \mathsf{H}^1(X(N), \mathcal{O}_{X(N)})^*,$$

where $\Omega_{X(N)}$ is the cotangent/canonical bundle of $X(N)$ and $\mathcal{O}_{X(N)}$ is its structure sheaf. The ***q*-expansion principle** says that

$$S_2(N;R) = S_2(N;\mathbb{Z}) \otimes R$$

when N is a unit in R or R is flat over $\mathbb{Z}$. For an appropriate model of $X(N)$, we also have an identification of these groups with $\mathsf{H}^0(X(N)_R, \Omega_{X(N)_R})$. This makes it reasonable to define $S_2(N;R) := S_2(N;\mathbb{Z}) \otimes R$ for any ring R satisfying one of those conditions.

1.2.2. For GL_1 we will be interested instead in the cohomology $\mathsf{H}^0(X(N), \mathcal{O}_{X(N)})$. Since $X(N) = \operatorname{Spec} \mathbb{Q}(\zeta_N)$ is affine, this group is just $\mathbb{Q}(\zeta_N)$ itself. The space of GL_1 cusp forms of level N should then be

$$\mathsf{H}^0(X(N),\ \mathcal{O}_{X(N)}) \otimes_{\mathbb{Q}} \mathbb{C} = \mathbb{Q}(\zeta_N) \otimes_{\mathbb{Q}} \mathbb{C}.$$

Remark 1.2.3. The group $\mathsf{H}^0(X(N),\ \mathcal{O}_{X(N)})$ is the same as the algebraic de Rham cohomology group $\mathsf{H}^0_{\mathrm{dR}}(X(N))$. The de Rham comparison theorem tells us that

$$\mathbb{Q}(\zeta_N) \otimes_{\mathbb{Q}} \mathbb{C} = \mathsf{H}^0_{\mathrm{dR}}(X(N)) \otimes_{\mathbb{Q}} \mathbb{C} \cong \mathsf{H}^0_{\mathrm{sing}}(X(N)(\mathbb{C}),\ \mathbb{C}) = \bigoplus_{\mathbb{Q}(\zeta_N) \hookrightarrow \mathbb{C}} \mathbb{C}.$$

It's a useful exercise to work out what this isomorphism is doing! In the GL_2 setting, the identity is a bit more complicated, since $\mathsf{H}^1_{\mathrm{dR}}(X(N)) \otimes_{\mathbb{Q}} \mathbb{C}$ is a direct sum of the holomorphic cusp forms $S_2(N;\mathbb{C})$ and the antiholomorphic cusp forms $\overline{S}_2(N;\mathbb{C})$.

GL_2	GL_1
$X(N),\ Y(N),\ X_0(N),\ Y_0(N)$	$X(N) = Y(N) = \operatorname{Spec}\mathbb{Q}(\zeta_N)$
$Y(N)(\mathbb{C}) = \mathrm{GL}_2(\mathbb{Q})\backslash\mathcal{H}^{\pm}\times\mathrm{GL}_2(\mathbb{A}_f)/\widehat{\Gamma}(N)$	$Y(N)(\mathbb{C}) = \mathbb{Q}^{\times}\backslash\{\pm 1\}\times\mathbb{A}_f^{\times}/\widehat{\Gamma}(N) = (\mathbb{Z}/N)^{\times}$
$\mathsf{H}^0(X(N),\ \Omega_{X(N)}) = S_2(N,\mathbb{Q})$	$S_0(N;\ \mathbb{C}) = \mathsf{H}^0(X(N),\ \mathcal{O}_{X(N)}) = \mathbb{Q}(\zeta_N)$
$\mathsf{H}_1^{\text{ét}}(X_{N,\overline{\mathbb{Q}}};\ \mathbb{Q}_\ell)$	$\mathsf{H}_0^{\text{ét}}(X_{N,\overline{\mathbb{Q}}};\ \mathbb{Q}_\ell) = \bigoplus_{\mathbb{Q}(\zeta_N)\hookrightarrow\overline{\mathbb{Q}}}\mathbb{Q}_\ell$
$T_p,\ \langle p\rangle$	$\langle p\rangle = \mathrm{Frob}_p$

Table 1: A dictionary between GL_2 objects and GL_1 objects.

1.2.4. Since we will be working with ℓ-adic Galois representations, we will be more interested in the ℓ-adic comparison theorem than the Betti one: picking $\overline{\mathbb{Q}}\hookrightarrow\overline{\mathbb{Q}}_\ell$, we have

$$\mathbb{Q}(\zeta_N)\otimes_{\mathbb{Q}}\overline{\mathbb{Q}}_\ell = \mathsf{H}^0_{\mathrm{dR}}(X(N))\otimes_{\mathbb{Q}}\overline{\mathbb{Q}}_\ell \cong \mathsf{H}^0_{\text{ét}}(X(N)_{\overline{\mathbb{Q}}},\ \overline{\mathbb{Q}}_\ell) = \bigoplus_{\mathbb{Q}(\zeta_N)\hookrightarrow\overline{\mathbb{Q}}}\overline{\mathbb{Q}}_\ell.$$

In the GL_2 case the correct statement requires period rings. We do have as a corollary of it that $S_2(N;\mathbb{C}_\ell)$ is a subspace of $\mathsf{H}^1_{\text{ét}}(X(N)_{\overline{\mathbb{Q}}},\ \mathbb{C}_\ell)$, which is one entry point into proving the Eichler–Shimura correspondence – namely, it puts cusp forms into a vector space with a Galois action. To describe this more precisely, we will need Hecke operators.

1.3 Hecke operators and automorphic representations

1.3.1. We have previously discussed how to turn a modular form into an automorphic representation. Formally this is accomplished in the following way: consider the Betti cohomology groups $\mathsf{H}^1_{\mathrm{sing}}(X(N)(\mathbb{C}),\ \mathbb{C})$ as N varies. Each space comes with an alternating bilinear form from Poincare duality which identifies $S_2(N;\ \mathbb{C})$ with the dual of $\overline{S}_2(N;\ \mathbb{C})$. We can rearrange this to give an inner product on $S_2(N;\mathbb{C})$ and $\mathsf{H}^1_{\mathrm{sing}}(X(N),\ \mathbb{C})$, the **Petersson inner product**. If N divides N' then pullback along $X(N')\to X(N)$ realizes $\mathsf{H}^1(X(N))$ as a subspace of $\mathsf{H}^1(X(N'))$. Consider the colimit of Hilbert spaces

$$\mathsf{H}^1(X,\ \mathbb{C}) := \varinjlim_{N}\mathsf{H}^1_{\mathrm{sing}}(X(N)(\mathbb{C}),\ \mathbb{C}).$$

Because $\mathrm{GL}_2(\mathbb{A})$ acts on the "full Shimura variety" $\varprojlim_N X(N)$, it also acts on the Hilbert space. Each newform generates an irreducible automorphic representation of $\mathrm{GL}_2(\mathbb{A})$ in this space.

1.3.2. For each prime p there is an object called the called the **spherical Hecke algebra** at p, which is the set of compactly supported functions on $\mathrm{GL}_n(\mathbb{Z}_p)\backslash\mathrm{GL}_n(\mathbb{Q}_p)/\mathrm{GL}_n(\mathbb{Z}_p)$, equipped with the convolution product. For GL_2 this algebra is generated by elements T_p and $\langle p\rangle$. For GL_1 the spherical Hecke algebra is generated just by $\langle p\rangle$.

The action of $\mathrm{GL}_2(\mathbb{A})$ on $\mathsf{H}^1(X, \mathbb{C})$ in particular gives an action of $\mathrm{GL}_2(\mathbb{Q}_p)$ for each prime p. Elements of the spherical Hecke algebra at p are, in particular, compactly supported functions on $\mathrm{GL}_2(\mathbb{Q}_p)$, so it makes sense to act with them on $\mathsf{H}^1(X, \mathbb{C})$ (say, by integration). If $\mathrm{GL}_2(\mathbb{Z}_p)$ acts trivially on some vector f (for instance if f is in $S_2(N;\ \mathbb{C})$ for N coprime to p), then the action is particularly easy to describe, since it can be written as a finite sum of action by coset representatives. Generally, T_p won't preserve $S_2(N;\mathbb{C})$, but it will if p is coprime to N or if we use $\Gamma_1(N)$ level structure instead of $\Gamma(N)$. When it does preserve $S_2(N;\mathbb{C})$, the element T_p viewed as an endomorphism of $S_2(N;\mathbb{C})$ is called a **Hecke operator**.

1.3.3. Let's return to the GL_1 case and compute the Hilbert space and the action of $\mathbb{A}^\times$, using the building blocks

$$S_0(N;\ \mathbb{C}) \coloneqq \mathsf{H}^0_{\mathrm{sing}}(X(N)(\mathbb{C}),\ \mathbb{C}) = \bigoplus_{\mathbb{Q}(\zeta_N)\hookrightarrow\mathbb{C}} \mathbb{C}.$$

The direct sum here is orthogonal with respect to the inner product. The full Shimura variety in this case is

$$\varprojlim \operatorname{Spec} \mathbb{Q}(\zeta_N) = \operatorname{Spec} \mathbb{Q}(\zeta_\infty),$$

the maximal cyclotomic extension of $\mathbb{Q}$. We view this as a profinite scheme over $\mathbb{Q}$, so in particular (if we pick an embedding $\mathbb{Q}(\zeta_\infty) \hookrightarrow \mathbb{C}$) we have that

$$X(\mathbb{C}) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_\infty) \mid \mathbb{Q}) = \widehat{\mathbb{Z}}^\times$$

with the profinite topology. The corresponding colimit of cohomology groups is the Hilbert space

$$\mathsf{H}^0(X,\ \mathbb{C}) = \operatorname*{colim}_{\overrightarrow{N}} \bigoplus_{\mathbb{Q}(\zeta_N)\hookrightarrow\mathbb{C}} \mathbb{C} = \mathsf{L}^2(X(\mathbb{C}),\ \mathbb{C}) \cong \mathsf{L}^2(\widehat{\mathbb{Z}}^\times,\ \mathbb{C}),$$

using the Haar measure on $\widehat{\mathbb{Z}}^\times$. In this identification, the function $\mathbb{1}_{\iota:\mathbb{Q}(\zeta_N)\hookrightarrow\mathbb{C}}$ in $S_0(N;\mathbb{C})$ is sent to

$$\mathbb{1}_{\sigma\,\mathrm{Gal}(\mathbb{Q}(\zeta_\infty)|\mathbb{Q}(\zeta_N))}$$

in $\mathsf{L}^2(\widehat{\mathbb{Z}}^\times)$ if ι is the σ-twist of the restriction of the "base embedding" by which we are identifying $X(\mathbb{C})$ and $\widehat{\mathbb{Z}}^\times$.

Finally, the ideles $\mathbb{A}^\times$ act on $\mathsf{L}^2(\widehat{\mathbb{Z}}^\times,\ \mathbb{C})$ via the identification

$$\widehat{\mathbb{Z}}^\times = \mathbb{Q}^\times \backslash \{\pm 1\} \times \mathbb{A}_f^\times$$

(where the action of $\mathbb{R}^\times$ on $\{\pm 1\}$ is via the sign character).

1.3.4. In terms of the spherical Hecke algebra for $\mathrm{GL}_1(\mathbb{Q}_p)$, the element $\langle p \rangle$ is just the function

$$\mathbb{1}_{p\mathbb{Z}_p^\times} : \mathbb{Q}_p^\times \to \mathbb{C}.$$

The action of $\langle p \rangle$ on a form $f \in S_0(N;\mathbb{C})$, where p is coprime to N, is easy to compute, since $\mathbb{Z}_p^\times$ fixes these forms. In this case the action is just the action of $p \in \mathbb{Q}_p^\times$ on f, which sends f to the function $x \mapsto f(xp)$ in $\mathsf{L}^2(\mathbb{Q}^\times \backslash (\widehat{\mathbb{Z}} \otimes \mathbb{Q})^\times)$, or equivalently the function $x \mapsto f(xp^{-1})$ in

$\mathsf{L}^2((\mathbb{Z}/N)^\times)$. Hence its action on $S_0(N;\mathbb{C})$ sends $\mathbb{1}_\sigma$ to $\mathbb{1}_{\sigma p}$. In other words, the action is via the Frobenius element Frob_p of $\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})$. More generally, we would need to average the action across $p\mathbb{Z}_p^\times$. On oldforms which are in the image of $S_0(Np^{-v_p(N)};\mathbb{C}) \hookrightarrow S_0(N;\mathbb{C})$, this averaging is such that the net effect is to lower the level of f to $Np^{-v_p(N)}$, then apply Frobenius at p, then raise the level back to N. The diamond operator annihilates forms not in this image.

We can define

$$S_0(N;\, \mathbb{Z}) := \bigoplus_{\mathbb{Q}(\zeta_N)\hookrightarrow\mathbb{C}} \mathbb{Z} \subseteq S_0(N;\, \mathbb{C}).$$

Then $S_0(N;\mathbb{Z})$ is preserved by the diamond operators $\langle p \rangle$. The subalgebra of $\mathrm{End}_{\mathbb{Z}}(S_0(N;\mathbb{Z}))$ generated by $\langle p \rangle$ for p not dividing N is called the (level N) **Hecke algebra**, denoted $\mathbf{T}_N$.

1.3.5. The diamond operators $\langle p \rangle$ and $\langle p' \rangle$ give the same element of the Hecke algebra whenever $p \equiv p' \mod N$. These elements act by permuting the basis of $S_0(N;\mathbb{Z})$ via an element of $\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})$. Furthermore, by, say, Dirichlet's theorem, every element of the Galois group arises as some diamond operator. These elements give a $\mathbb{Z}$-basis for $\mathbf{T}_N$, so we conclude

$$\mathbf{T}_N = \mathbb{Z}[\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})] = \mathbb{Z}[(\mathbb{Z}/N)^\times],$$

the group ring of $(\mathbb{Z}/N)^\times$.

2 The statement of R = T

2.1 Properties of the Hecke algebra

2.1.1. The action of $\mathbf{T}_N$ on $S_0(N;\mathbb{C})$ decomposes it into a sum of one dimensional subspaces. We call a generator of one of these subspaces an **eigenform**. Pick an embedding $\mathbb{Q}(\zeta_\infty) \hookrightarrow \mathbb{C}$ so that we can identify $S_0(N;\mathbb{C})$ with $\mathsf{L}^2((\mathbb{Z}/N)^\times)$. Then we normalize our eigenforms so that their value at the identity is 1. Recall that $\langle p \rangle$ for p not dividing N will act via Frob_p on this space. By Dirichlet's theorem, these Frobenius elements generate $(\mathbb{Z}/N)^\times$, so the decomposition into Hecke eigenforms is exactly the decomposition into characters of $(\mathbb{Z}/N)^\times$. Since we only need diamond operators at all but finitely many primes to distinguish different eigenforms, this is a version of **strong multiplicity one**.

If N divides N', then pullback along the map $X(N') \to X(N)$ gives an inclusion $S_0(N;\, \mathbb{C}) \hookrightarrow S_0(N';\, \mathbb{C})$ which is also compatible with the Hecke action away from N'. Eigenforms of $S_0(N';\, \mathbb{C})$ which are not in the image of any of these maps for N strictly dividing N' are called **newforms**. The eigenforms which are not newforms are called **oldforms**.

For GL_2, strong multiplicity one only applies to newforms. To further distinguish newforms and oldforms, one needs to work with Hecke operators at primes that divide N.

2.1.2. Given an eigenform f, the Hecke eigenvalues of f generate a finite field extension $K_f \mid \mathbb{Q}$. The decomposition

$$S_0(N;\, \mathbb{C}) = \bigoplus_f \mathbb{C}f$$

induces an algebra decomposition

$$\mathbf{T}_N \otimes \mathbb{C} = \prod_f \mathbb{C}.$$

In other words, $S_0(N;\mathbb{C})$ is a rank 1 free module over $\mathbf{T}_N \otimes \mathbb{C}$. We can refine this further: in fact, $S_0(N;\mathbb{Z})$ is a rank 1 free module over $\mathbf{T}_N$. (The analogous statement over $\mathbb{C}$ is still true for GL_2, but $S_2(N;\mathbb{Z})$ is not generally free over $\mathbf{T}_N$.)

2.1.3. The ring $\mathbf{T}_N$ has Krull dimension 1; its prime ideals are either maximal or minimal. The minimal primes are in bijection with $G_{\mathbb{Q}}$-orbits of normalized eigenforms in $S_0(N;\mathbb{C})$. Minimal primes also biject to $G_{\mathbb{Q}_\ell}$-orbits of normalized eigenforms in $S_0(N;\overline{\mathbb{Q}}_\ell)$ for any ℓ. In either case the prime associated to an eigenform f contains exactly the elements of the Hecke algebra which annihilate f.

Maximal primes of $\mathbf{T}_N$ correspond to eigenforms over fields of finite characteristic. More precisely, primes of $\mathbf{T}_N$ which lie over the prime $(\ell) \subset \mathbb{Z}$ biject to $G_{\mathbb{F}_\ell}$-orbits of eigenforms in $S_0(N;\overline{\mathbb{F}}_\ell)$. If a minimal prime corresponds to (the orbit of) an eigenform $f \in S_0(N;\overline{\mathbb{Q}}_\ell)$, then the maximal primes over it correspond to (the orbits of) the possible reductions of f in $S_0(N;\ \overline{\mathbb{F}}_\ell)$. For GL_1 such a reduction is unique: there is exactly one maximal prime over each minimal prime. For GL_2 this is not the case; minimal primes correspond to newforms, and a newform can have several associated oldforms that have different reductions. In either case, the characters of the localization of $\mathbf{T}_N$ at a maximal ideal correspond to eigenforms with a given reduction.

2.2 Eichler–Shimura

2.2.1. The analog of the Eichler–Shimura relation for GL_1 will feel somewhat tautological thanks to our explicit descriptions in Section 2.1.1. Given an eigenform $f \in S_0(N;\ \mathbb{C})$, we would like a Galois representation ρ_f such that the trace of Frobenius can be computed in terms of f.

There are a couple of ways to do this. One way is to use a comparison theorem to realize f as an element of $\mathsf{H}^0_{\text{ét}}(X(N)_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell)$ and take the Galois representation it generates. (This is dual to the Galois representation usually constructed in a Tate module.) Instead, we will leverage the fact that $\mathsf{H}_0^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\mathbb{Q}_\ell)$ is free of rank 1 over $\mathbf{T}_N \otimes \mathbb{Q}_\ell$. Write K_f for the number field generated by the Hecke eigenvalues of f. Pick a place λ of K_f dividing ℓ and write K_λ for the associated local field. Then the eigenform f determines a map

$$\mathbf{T}_{N,\mathbb{Q}_\ell} := \mathbf{T}_N \otimes \mathbb{Q}_\ell \to K_\lambda.$$

The Galois representation associated to f is

$$\mathsf{H}_0^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\ \mathbb{Q}_\ell) \otimes_{\mathbf{T}_{N,\mathbb{Q}_\ell}} K_\lambda.$$

Since the original module was rank 1 over $\mathbf{T}_{N,\mathbb{Q}_\ell}$, this is a one-dimensional Galois representation over K_λ. The Hecke algebra acts on étale homology in a manner dual to its action on S_0. Hence the Eichler–Shimura relation is the following.

Proposition 2.2.2. *If p is coprime to N, then p is unramified in ρ_f and $\rho_f(\mathrm{Frob}_p^{-1}) = \langle p \rangle$.*

2.2.3. We can compute various things regarding ρ_f. Most importantly, the conductor of ρ_f is N when f is a newform in $S_0(N;\mathbb{C})$. Furthermore, ρ_f is always an **Artin character**, meaning it has finite image. Equivalently, it has Hodge–Tate weight 0. Our goal will be to prove the following version of the Modularity Theorem.

Theorem 2.2.4. *The association $f \mapsto \rho_f$ gives a bijection between newforms of level N and Artin characters of conductor N.*

2.2.5. The theorem of Wiles and Taylor–Wiles, combined with later work of Breuil, Conrad, and Diamond, is analogous. They show that $f \mapsto \rho_f$ gives a bijection between newforms of level $\Gamma_0(N)$ with rational Hecke eigenvalues, and the Galois representations of the form $\mathrm{H}_1^{\text{ét}}(E_{\overline{\mathbb{Q}}};\, \mathbb{Q}_\ell)$ for an elliptic curve $E/\mathbb{Q}$ of conductor N.

2.3 The Galois deformation ring R

2.3.1. We have seen that modular representations arising from level N newforms are Artin characters with conductor N. In order to prove Theorem 2.2.4, we will consider the category of all Artin characters using various coefficients. As in last week's talk, the functor sending a ring A to the Artin characters with coefficients in A is a pro-representable functor. It can be pro-represented by the topological ring $\mathbf{R}^\square = \mathbb{Z}[G_{\mathbb{Q}}^{ab}]$, which is the group ring of the profinite group $G_{\mathbb{Q}}^{ab} = \mathrm{Gal}(\mathbb{Q}^{ab} \mid \mathbb{Q})$. Because we are working with characters, there is no difference between framed deformations and unframed deformations, so $\mathbf{R}^\square = \mathbf{R}$.

2.3.2. The condition of an Artin character having conductor dividing N is a deformation condition. We let $\mathbf{R}_N$ denote the corresponding pro-representing ring. A map

$$\mathbf{R} = \mathbb{Z}[G_{\mathbb{Q}}^{ab}] \to A$$

gives a $G_{\mathbb{Q}}$ representation in A of conductor dividing N if and only if each ramification group $G_p^{v_p(N)}$ acts trivially. So we can explicitly describe $\mathbf{R}_N$ as $\mathbb{Z}[G_{\mathbb{Q},N}^{ab}]$, where $G_{\mathbb{Q},N}^{ab}$ is the quotient of $G_{\mathbb{Q}}^{ab}$ by all these ramification groups that should act trivially. This is the same as the Galois group of $\mathbb{Q}_N \mid \mathbb{Q}$, where $\mathbb{Q}_N$ is the maximal abelian extension of conductor N. We remark that $\mathbf{R}_N$ is Noetherian, whereas $\mathbf{R}$ is not.

2.3.3. Recall that the maximal ideals of the Hecke algebra $\mathbf{T}_N$ over $(\ell) \subset \mathbb{Z}$ corresponded to reductions of the associated $\mathbb{Q}_\ell$-newform. The $\mathbb{Q}_\ell$-newforms themselves correspond to minimal primes of $\mathbf{T}$. A similar story is true on the Galois side.

The ring $\mathbf{R}_N$ has Krull dimension 1, and its minimal primes and its maximal primes over (ℓ) correspond to Galois orbits of Artin characters $G_{\mathbb{Q}} \to \overline{\mathbb{Q}}_\ell^\times$ and $G_{\mathbb{Q}} \to \overline{\mathbb{F}}_\ell^\times$, respectively. A minimal prime is contained in a maximal prime if the associated $\overline{\mathbb{Q}}_\ell$ representation reduces to the associated $\overline{\mathbb{F}}_\ell$ representation.

Remark 2.3.4. It would be more accurate to call $\mathbf{R}_N$ a universal representation ring rather than a deformation ring, since we have not yet fixed a residual representation which we would like to deform. Such a choice corresponds to a maximal ideal of $\mathbf{R}_N$, and the associated deformation ring is the completion of $\mathbf{R}_N$ at that ideal.

2.4 The map from R to T

2.4.1. Given an eigenform f in $S_0(N;\mathbb{C})$, we discussed how to associate a Galois representation ρ_f with coefficients in an ℓ-adic field K_λ. We will want to construct a map from $\mathbf{R}$ to $\mathbf{T}$, so we will need to understand the association $f \mapsto \rho_f$ for eigenforms with coefficients in more general rings.

We can think of a map from the Hecke algebra $\mathbf{T}_N$ to a ring A as a compatible assignment of Hecke eigenvalues in A to each diamond operator $\langle p \rangle$ for $p \nmid N$. If the ring A is a domain, then this uniquely determines an eigenform in $S_0(N;A)$, but we will also allow for non-domains. To repeat the construction of ρ_f in this context, we restrict to rings which are $\mathbb{Z}_\ell$-algebras. In this case, a map $\mathbf{T}_N \to A$ extends to

$$\mathbf{T}_{N,\mathbb{Z}_\ell} := \mathbf{T}_N \otimes \mathbb{Z}_\ell \to A.$$

Then we get a Galois representation with coefficients in A via

$$\mathsf{H}_0^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\, \mathbb{Z}_\ell) \otimes_{\mathbf{T}_{N,\mathbb{Z}_\ell}} A.$$

This is free of rank 1 over A, since $\mathsf{H}_0^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\mathbb{Z}_\ell)$ is free of rank 1 over $\mathbf{T}_{N,\mathbb{Z}_\ell}$. Specializing to the case $A = \mathbf{T}_{N,\mathbb{Z}_\ell}$, we get a map

$$G_{\mathbb{Q}} \to \operatorname{End}_{\mathbf{T}_{N,\mathbb{Z}_\ell}}(\mathsf{H}_0^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\, \mathbb{Z}_\ell)) = \mathbf{T}_{N,\mathbb{Z}_\ell}.$$

This is itself an Artin character of conductor N, so it is classified by a map $\mathbf{R}_{N,\mathbb{Z}_\ell} \to \mathbf{T}_{N,\mathbb{Z}_\ell}$.

Remark 2.4.2. This construction does not work as written for GL_2. There are two Hecke algebras at play in that case: the "full" Hecke algebra generated by all Hecke and diamond operators, and the "anemic" or "reduced" Hecke algebra generated by only the operators at primes not dividing N. The anemic Hecke algebra is the one appearing in the $\mathbf{R} = \mathbf{T}$ theorem, but the full Hecke algebra has nicer interaction with modular forms. For instance, $\mathsf{H}_1^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\mathbb{Q}_\ell)$ is free of rank 2 over the full Hecke algebra $\mathbf{T}_{N,\mathbb{Q}_\ell}$. Unlike for GL_1, however, $\mathsf{H}_1^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\mathbb{Z}_\ell)$ is not (generally) free over either ring. It turns out there are maximal ideals of $\mathbf{T}$ such that, after localization, the anemic and full Hecke algebras are isomorphic and $\mathsf{H}_1^{\text{ét}}(X(N)_{\overline{\mathbb{Q}}};\mathbb{Z}_\ell)$ is free. This takes some work to show. Instead, the map $\mathbf{R} \to \mathbf{T}$ is usually constructed explicitly, since the anemic Hecke algebra has an explicit description that we will see in a moment. The anemic and full Hecke algebras coincide for GL_1.

Remark 2.4.3. Using our descriptions of $\mathbf{R}_N$ and $\mathbf{T}_N$ as the group rings of $G_{\mathbb{Q},N}^{ab}$ and $\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})$, respectively, we can even write the map $\mathbf{R}_N \to \mathbf{T}_N$ integrally. Since $\mathbb{Q}(\zeta_N)$ has conductor N, its Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_N) \mid \mathbb{Q})$ is a quotient of the group $G_{\mathbb{Q},N}^{ab}$. The induced map on group rings gives a surjection $\mathbf{R}_N \twoheadrightarrow \mathbf{T}_N$.

2.4.4. Here we give a different description of $\mathbf{T}_N$, which has an analog for GL_2. If f is an eigenform in $S_0(N;\mathbb{C})$, then let $\mathcal{O}_f$ be the subring of $\mathbb{C}$ generated by the Hecke eigenvalues of f. The ring $\mathcal{O}_f$ is an order in its fraction field K_f. Then we write

$$\widetilde{\mathbf{T}}_N = \prod_f \mathcal{O}_f,$$

where the product is over newforms of level dividing N. Then the map

$$\mathbf{T}_N \to \widetilde{\mathbf{T}}_N,$$

sending a diamond operator to its eigenvalue on each eigenform, is an injection. We can then describe a map from $\mathbf{R}_N$ to $\widetilde{\mathbf{T}}_N$ by compiling the maps $\mathbf{R}_N \to \mathcal{O}_f$ that are induced by ρ_f, for each f. The image of this map is $\mathbf{T}_N$.

2.4.5. We have constructed (in several ways) a map $\mathbf{R}_N \to \mathbf{T}_N$. If this map were an isomorphism, then any Artin character of level N, classified by a map $\mathbf{R}_N \to \mathbb{C}$, would also come with a Hecke eigenform, classified by a map $\mathbf{T}_N \to \mathbb{C}$. In other words, if $\mathbf{R}_N \to \mathbf{T}_N$ is an isomorphism, then every level N Artin character is ρ_f for some eigenform f. Hence to prove Theorem 2.2.4, it is enough to show

Theorem 2.4.6. *The map $\mathbf{R}_N \to \mathbf{T}_N$ is an isomorphism.*

2.4.7. To keep with the analogy to the GL_2 story, what we will actually do is show that $\mathbf{R}_N \to \mathbf{T}_N$ becomes an isomorphism after localization at certain maximal ideals of $\mathbf{R}_N$. Because a surjective map of Noetherian local rings is an isomorphism after completion if and only if it is an isomorphism, it is enough to check that the maps $\mathbf{R}_{N,\mathfrak{m}} \to \mathbf{T}_{N,\mathfrak{m}}$ of complete local rings are isomorphisms. If we complete at a maximal ideal $\mathfrak{m}_{\overline{\rho}}$ which is in the preimage of a maximal ideal of $\mathbf{T}_N$, corresponding to a modular Galois character $\overline{\rho}$ over $\overline{\mathbb{F}}_\ell$, then the statement that $\mathbf{R}_{N,\mathfrak{m}_{\overline{\rho}}} \to \mathbf{T}_{N,\mathfrak{m}_{\overline{\rho}}}$ is an isomorphism implies that all $\overline{\mathbb{Q}}_\ell$-representations lifting $\overline{\rho}$ are modular. Whereas Theorem 2.4.6 directly implies the Modularity Theorem, we will focus on an *a priori* weaker version, the **Modularity Lifting Theorem**. In the language of $\mathbf{R}$ and $\mathbf{T}$, this says

Theorem 2.4.8. *Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{R}_N$ which is the preimage of a maximal ideal of $\mathbf{T}_N$. Then $\mathbf{R}_{N,\mathfrak{m}} \to \mathbf{T}_{N,\mathfrak{m}}$ is an isomorphism.*