

9.5 Galois representations and modular forms

This section associates Galois representations to modular curves and then decomposes them into 2-dimensional representations associated to modular forms.

Let N be a positive integer and let ℓ be prime. The modular curve $X_1(N)$ is a projective nonsingular algebraic curve over $\mathbf{Q}$. Let g denote its genus. The curve $X_1(N)_{\mathbf{C}}$ over $\mathbf{C}$ defined by the same equations can also be viewed as a compact Riemann surface. As in Chapter 6 the Jacobian of the modular curve is a g -dimensional complex torus obtained from integration modulo homology,

$$J_1(N) = \text{Jac}(X_1(N)_{\mathbf{C}}) = \mathcal{S}_2(\Gamma_1(N))^{\wedge} / H_1(X_1(N)_{\mathbf{C}}, \mathbf{Z}) \cong \mathbf{C}^g / \Lambda_g.$$

The Picard group of the modular curve is the Abelian group of divisor classes on the points of $X_1(N)$,

$$\text{Pic}^0(X_1(N)) = \text{Div}^0(X_1(N)) / \text{Div}^{\ell}(X_1(N)).$$

By the methods of Section 7.9, $\text{Pic}^0(X_1(N))$ can be identified with a subgroup of $\text{Pic}^0(X_1(N)_{\mathbf{C}})$, and the complex Picard group is naturally isomorphic to the Jacobian by Abel's Theorem as in Section 6.1. Thus there is an inclusion of ℓ^n -torsion,

$$i_n : \text{Pic}^0(X_1(N))[\ell^n] \longrightarrow \text{Pic}^0(X_1(N)_{\mathbf{C}})[\ell^n] \cong (\mathbf{Z}/\ell^n \mathbf{Z})^{2g}.$$

Recall that Igusa's Theorem (Theorem 8.6.1) states that $X_1(N)$ has good reduction at primes $p \nmid N$, so also there is a natural surjective reduction map $\text{Pic}^0(X_1(N)) \longrightarrow \text{Pic}^0(\tilde{X}_1(N))$ restricting to

$$\pi_n : \text{Pic}^0(X_1(N))[\ell^n] \longrightarrow \text{Pic}^0(\tilde{X}_1(N))[\ell^n].$$

We state without proof two generalizations of facts we have used about elliptic curves:

- The inclusion i_n is in fact an isomorphism.
- So is the surjection π_n for $p \nmid \ell N$.

These follow from results of algebraic geometry. Specifically, if a curve X over a field $\mathbf{k}$ has genus g and M is coprime to $\text{char}(\mathbf{k})$ then $\text{Pic}^0(X)[M] \cong (\mathbf{Z}/M\mathbf{Z})^{2g}$, and if a curve X over $\mathbf{Q}$ has good reduction at a prime $p \nmid M$ then the reduction map is injective on $\text{Pic}^0(X)[M]$.

The ℓ -adic Tate module of $X_1(N)$ is

$$\text{Ta}_{\ell}(\text{Pic}^0(X_1(N))) = \varprojlim_n \{\text{Pic}^0(X_1(N))[\ell^n]\}.$$

Similarly to the previous section, choosing bases of $\text{Pic}^0(X_1(N))[\ell^n]$ compatibly for all n shows that

$$\mathrm{Ta}_\ell(\mathrm{Pic}^0(X_1(N))) \cong \mathbf{Z}_\ell^{2g}.$$

Any automorphism $\sigma \in G_{\mathbf{Q}}$ defines an automorphism of $\mathrm{Div}^0(X_1(N))$,

$$\left(\sum n_P(P) \right)^\sigma = \sum n_P(P^\sigma).$$

Since $\mathrm{div}(f)^\sigma = \mathrm{div}(f^\sigma)$ for any $f \in \overline{\mathbf{Q}}(X_1(N))$, the automorphism descends to $\mathrm{Pic}^0(X_1(N))$,

$$\mathrm{Pic}^0(X_1(N)) \times G_{\mathbf{Q}} \longrightarrow \mathrm{Pic}^0(X_1(N)). \quad (9.13)$$

The field extension $\mathbf{Q}(\mathrm{Pic}^0(X_1(N))[\ell^n])/\mathbf{Q}$ is Galois for each $n \in \mathbf{Z}^+$, so the action restricts to $\mathrm{Pic}^0(X_1(N))[\ell^n]$. For each n there is a commutative diagram

$$\begin{array}{ccc} & G_{\mathbf{Q}} & \\ \swarrow & & \searrow \\ \mathrm{Aut}(\mathrm{Pic}^0(X_1(N))[\ell^n]) & \longleftarrow & \mathrm{Aut}(\mathrm{Pic}^0(X_1(N))[\ell^{n+1}]). \end{array}$$

Again as in the previous section this leads to a continuous homomorphism

$$\rho_{X_1(N),\ell} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_{2g}(\mathbf{Z}_\ell) \subset \mathrm{GL}_{2g}(\mathbf{Q}_\ell).$$

This is the $2g$ -dimensional Galois representation associated to $X_1(N)$.

Recall from Chapter 6 that the Hecke algebra over $\mathbf{Z}$ is the algebra of endomorphisms of $\mathcal{S}_2(\Gamma_1(N))$ generated over $\mathbf{Z}$ by the Hecke operators,

$$\mathbb{T}_{\mathbf{Z}} = \mathbf{Z}[\{T_n, \langle n \rangle : n \in \mathbf{Z}^+\}].$$

The Hecke algebra acts on $\mathrm{Pic}^0(X_1(N))$, cf. the bottom rows of diagrams (7.18) and (7.19),

$$\mathbb{T}_{\mathbf{Z}} \times \mathrm{Pic}^0(X_1(N)) \longrightarrow \mathrm{Pic}^0(X_1(N)). \quad (9.14)$$

Since the action is linear it restricts to ℓ -power torsion, and so it extends to $\mathrm{Ta}_\ell(\mathrm{Pic}^0(X_1(N)))$. From Section 7.9 the Hecke action is defined over $\mathbf{Q}$. So the Galois action (9.13) and the Hecke action (9.14) on $\mathrm{Pic}^0(X_1(N))$ commute, and therefore so do the two actions on $\mathrm{Ta}_\ell(\mathrm{Pic}^0(X_1(N)))$.

Theorem 9.5.1. *Let ℓ be prime and let N be a positive integer. The Galois representation $\rho_{X_1(N),\ell}$ is unramified at every prime $p \nmid \ell N$. For any such p let $\mathfrak{p} \subset \overline{\mathbf{Z}}$ be any maximal ideal over p . Then $\rho_{X_1(N),\ell}(\mathrm{Frob}_{\mathfrak{p}})$ satisfies the polynomial equation*

$$x^2 - T_p x + \langle p \rangle p = 0.$$

Similarly to how we have abbreviated $T_{p,*}$ to T_p on $\text{Pic}^0(X_1(N))$ all along, the last formula in the theorem omits the asterisk from the subscript of both Hecke operators on $\text{Ta}_\ell(\text{Pic}^0(X_1(N)))$. As in the previous section, the vector space $V_\ell(X_1(N)) = \text{Ta}_\ell(\text{Pic}^0(X_1(N))) \otimes \mathbf{Q}$ can be taken as the Galois representation rather than $\rho_{X_1(N),\ell}$, there is a corresponding commutative diagram, and the theorem can be rephrased appropriately.

Proof. Let $p \nmid \ell N$ and let $\mathfrak{p}$ lie over p . As in the proof of Theorem 9.4.1 there is a commutative diagram

$$\begin{array}{ccc} D_{\mathfrak{p}} & \longrightarrow & \text{Aut}(\text{Pic}^0(X_1(N))[\ell^n]) \\ \downarrow & & \downarrow \\ G_{\mathbf{F}_p} & \longrightarrow & \text{Aut}(\text{Pic}^0(\tilde{X}_1(N))[\ell^n]). \end{array}$$

The map down the right side is an isomorphism as explained at the beginning of the section. Similarly to before, $I_{\mathfrak{p}} \subset \ker \rho_{X_1(N),\ell}$.

For the second part, the Eichler–Shimura Relation (Theorem 8.7.2) restricts to ℓ^n -torsion,

$$\begin{array}{ccc} \text{Pic}^0(X_1(N))[\ell^n] & \xrightarrow{T_p} & \text{Pic}^0(X_1(N))[\ell^n] \\ \downarrow & & \downarrow \\ \text{Pic}^0(\tilde{X}_1(N))[\ell^n] & \xrightarrow{\sigma_{p,*} + \langle p \rangle_* \sigma_p^*} & \text{Pic}^0(\tilde{X}_1(N))[\ell^n]. \end{array}$$

The same diagram but with $\text{Frob}_{\mathfrak{p}} + \langle p \rangle p \text{Frob}_{\mathfrak{p}}^{-1}$ across the top row instead also commutes, cf. (8.15). Since the vertical arrows are isomorphisms, $T_p = \text{Frob}_{\mathfrak{p}} + \langle p \rangle p \text{Frob}_{\mathfrak{p}}^{-1}$ on $\text{Pic}^0(X_1(N))[\ell^n]$. This holds for all n , so the equality extends to $\text{Ta}_\ell(\text{Pic}^0(\tilde{X}_1(N)))$. The result follows. $\square$

To proceed from Picard groups to modular forms, consider a normalized eigenform

$$f \in \mathcal{S}_2(N, \chi).$$

Recall from Chapter 6 that the Hecke algebra contains an ideal associated to f , the kernel of the eigenvalue map,

$$I_f = \{T \in \mathbb{T}_{\mathbf{Z}} : Tf = 0\},$$

and the Abelian variety of f is defined as

$$A_f = J_1(N)/I_f J_1(N).$$

This is a complex analytic object. We do not define an algebraic version of it because its role here is auxiliary. By (6.12) and Exercise 6.5.2 there is an isomorphism

$$\mathbb{T}_{\mathbf{Z}}/I_f \xrightarrow{\sim} \mathcal{O}_f \quad \text{where} \quad \mathcal{O}_f = \mathbf{Z}[\{a_n(f) : n \in \mathbf{Z}^+\}].$$

Under this isomorphism each Fourier coefficient $a_p(f)$ acts on A_f as $T_p + I_f$. Also, $\mathcal{O}_f$ contains the values $\chi(n)$ for $n \in \mathbf{Z}^+$ and $\chi(p)$ acts on A_f as $\langle p \rangle + I_f$. The ring $\mathcal{O}_f$ generates the number field of f , denoted $\mathbf{K}_f$. The extension degree $d = [\mathbf{K}_f : \mathbf{Q}]$ is also the dimension of A_f as a complex torus. As with elliptic curves and modular curves, the Abelian variety has an ℓ -adic Tate module,

$$\mathrm{Ta}_\ell(A_f) = \varprojlim_n \{A_f[\ell^n]\} \cong \mathbf{Z}_\ell^{2d}.$$

The action of $\mathcal{O}_f$ on A_f is defined on ℓ -power torsion and thus extends to an action on $\mathrm{Ta}_\ell(A_f)$. The following lemma shows that $G_{\mathbf{Q}}$ acts on $\mathrm{Ta}_\ell(A_f)$ as well.

Lemma 9.5.2. *The map $\mathrm{Pic}^0(X_1(N))[\ell^n] \longrightarrow A_f[\ell^n]$ is a surjection. Its kernel is stable under $G_{\mathbf{Q}}$.*

Proof. Multiplication by ℓ^n is surjective on the complex torus $J_1(N)$. This makes it surjective on $I_f J_1(N)$ as well. Indeed, any $y \in I_f J_1(N)$ takes the form $y = \sum_i T_i y_i$ with $T_i \in I_f$ and $y_i \in J_1(N) = \ell^n J_1(N)$ for each i , so $y = \sum_i T_i (\ell^n x_i) = \ell^n \sum_i T_i x_i \in \ell^n I_f J_1(N)$ as desired.

To show the first statement of the lemma, take any $y \in A_f[\ell^n]$. Then $y = x + I_f J_1(N)$ for some $x \in J_1(N)$ such that $\ell^n x \in I_f J_1(N)$. Thus $\ell^n x = \ell^n x'$ for some $x' \in I_f J_1(N)$ by the previous paragraph. The difference $x - x'$ lies in $J_1(N)[\ell^n] = \mathrm{Pic}^0(X_1(N))[\ell^n]$ and maps to y as desired.

The kernel is $\mathrm{Pic}^0(X_1(N))[\ell^n] \cap I_f J_1(N) = (I_f J_1(N))[\ell^n]$. We claim that the inclusion $(I_f \mathrm{Pic}^0(X_1(N)))[\ell^n] \subset (I_f J_1(N))[\ell^n]$ is in fact an equality. To see this, let $\mathcal{S}_2 = \mathcal{S}_2(I_1(N))$ and $H_1 = H_1(X_1(N)_{\mathbf{C}}, \mathbf{Z}) \subset \mathcal{S}_2^\wedge$. Thus $J_1(N) = \mathcal{S}_2^\wedge/H_1$ and

$$I_f J_1(N) = (I_f \mathcal{S}_2^\wedge + H_1)/H_1 \cong I_f \mathcal{S}_2^\wedge / (H_1 \cap I_f \mathcal{S}_2^\wedge). \quad (9.15)$$

Proposition 6.2.4 shows that $I_f H_1$ is a subgroup of $H_1 \cap I_f \mathcal{S}_2^\wedge$ with some finite index M . This shows that $M(H_1 \cap I_f \mathcal{S}_2^\wedge) \subset I_f H_1$. Now suppose that $y \in (I_f J_1(N))[\ell^n]$. Then (9.15) shows that $y = x + H_1 \cap I_f \mathcal{S}_2^\wedge$ with $x \in I_f \mathcal{S}_2^\wedge$, and since $\ell^n y = 0$ this implies $\ell^n x \in H_1 \cap I_f \mathcal{S}_2^\wedge$. Therefore $M \ell^n x \in M(H_1 \cap I_f \mathcal{S}_2^\wedge) \subset I_f H_1$, and so $x \in I_f (M^{-1} \ell^{-n} H_1)$. It follows that $y \in I_f (J_1(N)[M \ell^n]) \subset I_f \mathrm{Pic}^0(X_1(N))$, and since $\ell^n y = 0$ the equality is proved. Thus the kernel is $(I_f (\mathrm{Pic}^0(X_1(N))))[\ell^n]$. This is stable under $G_{\mathbf{Q}}$ as desired since the Galois and Hecke actions on $\mathrm{Pic}^0(X_1(N))$ commute. $\square$

So $G_{\mathbf{Q}}$ acts on $A_f[\ell^n]$ and therefore on $\mathrm{Ta}_\ell(A_f)$. The action commutes with the action of $\mathcal{O}_f$ since the $G_{\mathbf{Q}}$ -action and the $\mathbb{T}_{\mathbf{Z}}$ -action commute on $\mathrm{Ta}_\ell(\mathrm{Pic}^0(X_1(N)))$. Choosing coordinates appropriately gives a Galois representation

$$\rho_{A_f, \ell} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_{2d}(\mathbf{Q}_\ell).$$

This is continuous because $\rho_{X_1(N),\ell}$ is continuous and (Exercise 9.5.1)

$$\rho_{X_1(N),\ell}^{-1}(U(n,g)) \subset \rho_{A_f,\ell}^{-1}(U(n,d)), \quad (9.16)$$

where $U(n,g) = \ker(\mathrm{GL}_{2g}(\mathbf{Z}_\ell) \rightarrow \mathrm{GL}_{2g}(\mathbf{Z}/\ell^n\mathbf{Z}))$ and similarly for $U(n,d)$. The representation is unramified at all primes $p \nmid \ell N$ since its kernel contains $\ker \rho_{X_1(N),\ell}$. For any such p let $\mathfrak{p} \subset \overline{\mathbf{Z}}$ be any maximal ideal over p . At the level of Abelian varieties, since T_p acts as $a_p(f)$ and $\langle p \rangle$ acts as $\chi(p)$, $\rho_{A_f,\ell}(\mathrm{Frob}_{\mathfrak{p}})$ satisfies the polynomial equation

$$x^2 - a_p(f)x + \chi(p)p = 0.$$

The Tate module $\mathrm{Ta}_\ell(A_f)$ has rank $2d$ over $\mathbf{Z}_\ell$. Since it is an $\mathcal{O}_f$ -module the tensor product $V_\ell(A_f) = \mathrm{Ta}_\ell(A_f) \otimes \mathbf{Q}$ is a module over $\mathcal{O}_f \otimes \mathbf{Q}_\ell = \mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell$.

Lemma 9.5.3. *$V_\ell(A_f)$ is a free module of rank 2 over $\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell$.*

Proof. Again let $\mathcal{S}_2 = \mathcal{S}_2(\Gamma_1(N))$ and $H_1 = H_1(X_1(N)_{\mathbf{C}}, \mathbf{Z}) \subset \mathcal{S}_2^\wedge$. Consider the quotients $\overline{\mathcal{S}_2^\wedge} = \mathcal{S}_2^\wedge / I_f \mathcal{S}_2^\wedge$ and $\overline{H}_1 = (H_1 + I_f \mathcal{S}_2^\wedge) / I_f \mathcal{S}_2^\wedge$. Then $A_f = \mathcal{S}_2^\wedge / (H_1 + I_f \mathcal{S}_2^\wedge) = \overline{\mathcal{S}_2^\wedge} / \overline{H}_1$. Thus $\overline{H}_1$ is an $\mathcal{O}_f$ -module whose $\mathbf{Z}$ -rank is $2d$. Since $\mathbf{K}_f$ is a field, $\overline{H}_1 \otimes \mathbf{Q}$ is a free $\mathbf{K}_f$ -module of rank 2, and therefore $\overline{H}_1 \otimes \mathbf{Q}_\ell = \overline{H}_1 \otimes \mathbf{Q} \otimes_{\mathbf{Q}} \mathbf{Q}_\ell$ is free of rank 2 over the ring $\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell$.

The $\mathcal{O}_f$ -linear isomorphisms $\ell^{-n} \overline{H}_1 / \overline{H}_1 \rightarrow \overline{H}_1 / \ell^n \overline{H}_1$ induced by multiplication by ℓ^n on $\ell^{-n} \overline{H}_1$ assemble to give an isomorphism of $\mathcal{O}_f \otimes \mathbf{Z}_\ell$ -modules,

$$\mathrm{Ta}_\ell(A_f) = \varprojlim_n \{A_f[\ell^n]\} = \varprojlim_n \{\ell^{-n} \overline{H}_1 / \overline{H}_1\} \cong \varprojlim_n \{\overline{H}_1 / \ell^n \overline{H}_1\} \cong \overline{H}_1 \otimes \mathbf{Z}_\ell,$$

where the transition maps in the last inverse limit are the natural projection maps. And now $V_\ell(A_f) = \mathrm{Ta}_\ell(A_f) \otimes \mathbf{Q} \cong \overline{H}_1 \otimes \mathbf{Z}_\ell \otimes \mathbf{Q} \cong \overline{H}_1 \otimes \mathbf{Q}_\ell$ is an isomorphism of modules over $\mathcal{O}_f \otimes \mathbf{Z}_\ell \otimes \mathbf{Q} = \mathbf{K}_f \otimes \mathbf{Q}_\ell$, showing that $V_\ell(A_f)$ is free. $\square$

The absolute Galois group $G_{\mathbf{Q}}$ acts $(\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell)$ -linearly on $V_\ell(A_f)$, and $V_\ell(A_f) \cong (\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell)^2$ by the lemma. Choose a basis B of $V_\ell(A_f)$ to get a homomorphism $G_{\mathbf{Q}} \rightarrow \mathrm{GL}_2(\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell)$. Also, (9.9) specializes to give $\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_\ell \cong \prod_{\lambda|\ell} \mathbf{K}_{f,\lambda}$, so for each λ we can compose the homomorphism with a projection to get

$$\rho_{f,\lambda} : G_{\mathbf{Q}} \rightarrow \mathrm{GL}_2(\mathbf{K}_{f,\lambda}).$$

This is continuous (Exercise 9.5.2(b)), making it a Galois representation. And $\ker(\rho_{A_f,\ell}) \subset \ker(\rho_{f,\lambda})$ (Exercise 9.5.2(c)). We have proved

Theorem 9.5.4. *Let $f \in \mathcal{S}_2(N, \chi)$ be a normalized eigenform with number field $\mathbf{K}_f$. Let ℓ be prime. For each maximal ideal λ of $\mathcal{O}_{\mathbf{K}_f}$ lying over ℓ there is a 2-dimensional Galois representation*

$$\rho_{f,\lambda} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\mathbf{K}_{f,\lambda}).$$

This representation is unramified at every prime $p \nmid \ell N$. For any such p let $\mathfrak{p} \subset \overline{\mathbf{Z}}$ be any maximal ideal lying over p . Then $\rho_{f,\lambda}(\mathrm{Frob}_{\mathfrak{p}})$ satisfies the polynomial equation

$$x^2 - a_p(f)x + \chi(p)p = 0.$$

In particular, if $f \in \mathcal{S}_2(\Gamma_0(N))$ then the relation is $x^2 - a_p(f)x + p = 0$.

Exercises

9.5.1. Establish (9.16). (A hint for this exercise is at the end of the book.)

9.5.2. (a) Let $i : \mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_{\ell} \longrightarrow \prod_{\lambda|\ell} \mathbf{K}_{f,\lambda}$ be the isomorphism of (9.9). For each λ , let e_{λ} be the element of $\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_{\ell}$ that is taken by i to $(0, \dots, 0, 1_{\mathbf{K}_{f,\lambda}}, 0, \dots, 0)$ and let $V_{f,\lambda} = e_{\lambda}(V_{\ell}(A_f))$. Show that each $V_{f,\lambda}$ is a 2-dimensional vector space over $\mathbf{K}_{f,\lambda}$ and that

$$V_{\ell}(A_f) = \bigoplus_{\lambda|\ell} V_{f,\lambda}.$$

Show that each $V_{f,\lambda}$ is invariant under the $G_{\mathbf{Q}}$ -action on $V_{\ell}(A_f)$. Show that if each $V_{f,\lambda}$ is given the basis $e_{\lambda}B$ over $\mathbf{K}_{f,\lambda}$ where B is the basis of $V_{\ell}(A_f)$ over $\mathbf{K}_f \otimes_{\mathbf{Q}} \mathbf{Q}_{\ell}$ in the section then each $\rho_{f,\lambda}$ is defined by the action of $G_{\mathbf{Q}}$ on $V_{f,\lambda}$. (Hints for this exercise are at the end of the book.)

(b) To show that $\rho_{f,\lambda}$ is continuous it suffices to show that the action

$$V_{f,\lambda} \times G_{\mathbf{Q}} \longrightarrow V_{f,\lambda}$$

is continuous. Explain why this statement is independent of whether $V_{f,\lambda}$ is viewed as a vector space over $\mathbf{K}_{f,\lambda}$ or over $\mathbf{Q}_{\ell}$. Explain why the action is continuous in the latter case.

(c) Use the decomposition from (a) to show that $\ker(\rho_{A_f,\ell}) \subset \ker(\rho_{f,\lambda})$ for each λ .

9.6 Galois representations and Modularity

This last section states the Modularity Theorem in terms of Galois representations, connects it to the arithmetic versions in Chapter 8, and describes how the modularity of elliptic curves is part of a broader conjecture. Finally we discuss how the modularity of Galois representations and of mod ℓ representations are related.

Definition 9.6.1. An irreducible Galois representation

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\mathbf{Q}_{\ell})$$

such that $\det \rho = \chi_{\ell}$ is **modular** if there exists a newform $f \in \mathcal{S}_2(\Gamma_0(M_f))$ such that $\mathbf{K}_{f,\lambda} = \mathbf{Q}_{\ell}$ for some maximal ideal λ of $\mathcal{O}_{\mathbf{K}_f}$ lying over ℓ and such that $\rho_{f,\lambda} \sim \rho$.