Deformation of a modular form and modular Galois representations

Haruzo Hida
UCLA, Los Angeles, CA 90095-1555, U.S.A.
UCLA student seminar, September 27, 2022.

*I would like to explain the motivation of my starting $p$-adic deformation theory of modular forms and modular Galois representations and the main points of the theory. All modular forms in this talk are elliptic either classical or $p$-adic. We write $X_?(N) (\? = 0,1)$ for compactified modular curves of the level $N$ groups $\Gamma_?(N) \subset SL_2(\mathbb{Z})$, and $Y_?(N) = X_?(N) - \{\text{cusps}\}$. Here are references:

[H86b] Inventiones 85 (1986) (Galois representation paper)
§0. A fundamental question in 1980. Consider a homogeneous space \( X := \text{GL}_2(\mathbb{Q}) \backslash \text{GL}_2(\mathbb{A})/\mathbb{A}^{	imes} \) with a Haar measure \( d\mu \) and an \( L^2 \) Hilbert space

\[
L^2_{\text{cusp}} := \{ f : X \to \mathbb{C} | \int_X |f|^2 d\mu < \infty, \int_{\mathbb{A}} f(\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} h) du = 0 \ \forall h \}. 
\]

By Harish-Chandra, the representation \( \Pi : \text{GL}_2(\mathbb{A}) \to \text{Aut}_\mathbb{C}(L^2_{\text{cusp}}) \) given by \( \Pi(f)(h) = f(hg) \) is a discrete direct sum of irreducible representations. Towards the end of 1980, I felt that having a discrete spectrum is unfortunate. One cannot easily come close to a harder irreducible constituent \( \pi \subset \Pi \) from another somehow simpler \( \pi_0 \) just by discreteness.

Question. Is there any other topology, keeping the information of \( \pi \) and \( \pi_0 \) to good extent, to have a continuous spectrum?
§1. Why not try $p$-adic topology? We need a $p$-adic replacement of $L^2_{\text{cusp}}$: $S(\mathbb{C}) = \text{GL}_2(\mathbb{Q}) \backslash \text{GL}_2(\mathbb{A}) / \mathbb{R}_+ \times \text{SO}_2(\mathbb{R})$ is the complex points of the Shimura curve (a pro-curve made of modular curves: $\Gamma \backslash \mathcal{H}$) defined over $\mathbb{Z}$. Consider a holomorphic cusp form $f: \mathcal{H} \to \mathbb{C}$ of weight $k$ on $\Gamma$. Writing $j \left( \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tau \right) = c\tau + d$, $f$ satisfies $f(\gamma(\tau)) = f(\tau)j(\gamma, \tau)^k$ for $\gamma \in \Gamma$. The automorphic factor $g \mapsto j(g, \tau)^k$ is a 1-cocycle having values in nowhere vanishing holomorphic functions on $\mathcal{H}$ as $j(gh, \tau) = j(g, h(\tau))j(g, \tau)$. Thus $\Gamma \backslash (\mathcal{H} \times \mathbb{C})$ by the action $(\tau, u) \mapsto (\gamma(\tau), j(\gamma, \tau)^k u)$ gives a cuspidal line bundles $\omega_{\text{cusp}}^k \subset \omega^k$ on $\Gamma \backslash \mathcal{H}$, which descends to $\mathbb{Z}$ canonically, as $\omega = \varphi_* \Omega^1_{\mathcal{E}/S}$ for the universal elliptic curve $\mathcal{E} \to S/\mathbb{Z}$. Similarly, writing $L(n; A) = AX^n + AX^{n-1}Y + \cdots + AY^n$ and letting $\text{SL}_2(\mathbb{Z})$ acts on $\mathcal{H} \times L(n; A)$ by $(\tau, P(X, Y)) \mapsto (\gamma(\tau), P((X, Y)^t \gamma^{-1}))$, we have the sheaf $\mathcal{L}(n; A)$ of locally constant sections from $\Gamma \backslash \mathcal{H}$ to $\Gamma \backslash (\mathcal{H} \times L(n; A))$. We replace $L^2$ by the $p$-adic completion of

\[
\begin{cases}
V_k := \lim_{\alpha} H^0(X_1(Np^\alpha) / \mathbb{Z}_p, \omega_{\text{cusp}}^k) & \text{coherent cohomology}, \\
TJ_n := \lim_{\alpha} H^{1}_{\text{et},!}(Y_1(Np^\alpha), \mathcal{L}(n; \mathbb{Z}_p)) & \text{pro-étale cohomology}.
\end{cases}
\]

Here $H_1^1 = \text{Im}(H^1_c \to H^1)$. 
§1. Why taking the limit for $Y_1(Np^n)$. Taking limit with respect to $p^n$ is an analogue from Kummer and Iwasawa, as $\mathbb{Z}_p^\times = \varprojlim_\alpha \Gamma_0(p^n) / \Gamma_1(p^n)$, the modules $V$ and $TJ_n$ are Iwasawa modules over $\mathbb{Z}_p[[\Gamma]] \cong \mathbb{Z}_p[[T]]$ for $\Gamma = \mathbb{Z}_p^\times / \{\text{torsion}\} = \gamma^\mathbb{Z}_p$. I wanted to make a $GL(2)$-version of Iwasawa theory. Let

$$
\hat{\Gamma}_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\hat{\mathbb{Z}}) \bigg| c \in N\hat{\mathbb{Z}} \right\},
$$
$$
\hat{\Gamma}_1(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \hat{\Gamma}_0(N) \bigg| a \equiv 1 \mod N\hat{\mathbb{Z}} \right\}.
$$

Then $Y?(N) = S / \hat{\Gamma}?(N) \cong \Gamma?(N) \backslash \mathfrak{H}$. For each $\pi \subset \Pi$, there is a minimal $N$ with $H^0(\hat{\Gamma}_1(N), \pi) = \mathbb{C}f_\pi$. If $f$ is holomorphic, we call $\pi$ is holomorphic, and $f = \sum_{n=1}^\infty a(n, f)e(n\tau)$. Decomposing

$$
\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}) \big| ad - bc = n, a \equiv 1 \mod N, N|c \} = \bigsqcup_i \alpha_i \Gamma_1(N),
$$

define the Hecke operator $T(n)$ by $f | T(n)(g) = \sum_i f(g \alpha_i^{(\infty)})$, we get $f | T(n) = a(n, f)f$ if $a(1, f) = 1$ as $H^0(\hat{\Gamma}_1(N), \pi) = \mathbb{C}f_\pi$, and $T(n)$ determines $f$ and hence $\pi$. Therefore, it is sufficient to take limit with respect to $\Gamma_1(Np^\alpha)$. 
§3. Weight reduction to the constant sheaf. Consider a morphism $I : L(n; \mathbb{Z}/p^\alpha \mathbb{Z}) \to \mathbb{Z}/p^\alpha \mathbb{Z}$ sending $P(X, Y) \mapsto P(1, 0)$. Note $\Gamma_1(p^\alpha) \ni \gamma \equiv \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \mod p^\alpha$. Thus $I(\gamma P(X, Y)) = P((1, 0) \begin{pmatrix} 1 & 0 \\ -u & 1 \end{pmatrix}) = I(P(X, Y))$, and hence $I : L(n; \mathbb{Z}/p^\alpha \mathbb{Z}) \to \mathbb{Z}/p^n \mathbb{Z}$ is an étale sheaf morphism over $Y_0(p^\alpha)$. Passing to the limit, we have a Galois and Hecke equivariant map: $TJ_n \to TJ_0$. We have well defined Hecke operators $T(n)$ and $T(l) = U(l)$ for $l | Np$ on $H^1_{et}$, and $I$ is also Hecke equivariant. Define for $k = n + 2$ and $\Lambda := \mathbb{Z}_p[[\Gamma]] = \mathbb{Z}_p[[T]]$ (by $\langle \gamma \rangle \mapsto 1 + T$ for the fixed generator $\gamma$ of $\Gamma$)

$$h_k(Np^\infty; \mathbb{Z}_p) = \Lambda[[T(n) : 0 < n \in \mathbb{Z}]] \subset \text{End}_\Lambda(TJ_n).$$

Independence Theorem: [EMI, Theorem 4.2.17]. We have a canonical isomorphism $i : h_k(Np^\infty; \mathbb{Z}_p) \cong h_2(Np^\infty; \mathbb{Z}_p)$ such that $i(T(n)) = T(n)$, $i(U(l)) = U(l)$ and $i(\langle z \rangle) = z^n \langle z \rangle$ for $z \in \mathbb{Z}_p^\times$, where $\langle z \rangle \in \lim_{\leftarrow} \Gamma_0(p^\alpha)/\Gamma_1(p^\alpha) \cong \mathbb{Z}_p^\times$ and $z^n \in \mathbb{Z}_p^\times \hookrightarrow \mathbb{Q}_p$.

Hereafter we just write $h := h_2(Np^\infty; \mathbb{Z}_p)$. 
§4. Finite level Hecke algebras. Define
\[ h_k(Np^\alpha; \mathbb{Z}_p) = \mathbb{Z}_p[T(n)|0 < n \in \mathbb{Z}] \subset \text{End}_{\mathbb{Q}_p}(H^1(Y_1(Np^\alpha), \mathcal{L}(n; \mathbb{Q}_p))). \]
Then by definition
\[ h = \lim_{\alpha} h_k(Np^\alpha; \mathbb{Z}_p) \text{ for all } 2 \leq k \in \mathbb{Z}. \]
By the Eichler–Shimura Hecke equivariant isomorphism:
\[ H^1(X_1(Np^\alpha), \mathcal{L}(n; \mathbb{C})) \cong H^0(X_1(Np^\alpha)/\mathbb{C}, \omega^k_{\text{cusp}}) \oplus H^0(X_1(Np^\alpha)/\mathbb{C}, \omega^k_{\text{cusp}}), \]
alternatively,
\[ h_k(Np^\alpha; \mathbb{Z}_p) := \mathbb{Z}_p[T(n)|0 < n \in \mathbb{Z}] \subset \text{End}_{\mathbb{Z}_p}(H^0(X_1(Np^\alpha)/\mathbb{Z}_p, \omega^k_{\text{cusp}})). \]
Thus \( h \) acts on \( V_k \).
§5. *q*-Expansion. The $J$-invariant has its $q$-expansion in $q^{-1}\mathbb{Z}[[q]]$ starting with $q^{-1}$; so, $J^{-1} = q + \cdots \in \mathbb{Z}[[q]]^\times$ gives rise to a parameter at $\infty$ of $X_0(1) = \mathbb{P}^1(J)/\mathbb{Z}$. Since $X_1(N)$ is étale over $\infty \in \mathbb{P}^1(J)$, $\hat{\mathcal{O}}_{X_1(N),\infty}/\mathbb{Z} = \mathbb{Z}[[q]]$. Thus $f \in H^0(X_1(Np^\alpha)/\mathbb{Z}_p, \omega^k_{\text{cus}})$ has $q$-expansion $f(q) = \sum_{n=1}^\infty a(n,f)q^n \in \mathbb{Z}_p[[q]]$. By the $q$-expansion principle, we have $V_k \hookrightarrow \mathbb{Z}_p[[q]]$ as $\mathbb{Z}_p[[q]]$ is $p$-adically complete.

Serre considered the space of $p$-adic modular forms as a $p$-adic completion of $\sum_k H^0(X_1(N)/\mathbb{Z}_p, \omega^k_{\text{cus}}) \subset \mathbb{Z}_p[[q]]$ in 1973. Soon after, Katz generalized this to the $p$-adic completion $V(Np^\alpha)$ of $\sum_k H^0(X_1(N)/\mathbb{Q}_p, \omega^k_{\text{cus}}) \cap \mathbb{Z}_p[[q]]$ inside $\mathbb{Z}_p[[q]]$, and via his notion of geometric modular forms, he remarked

$$V(Np^\alpha) = V(N) \text{ inside } \mathbb{Z}_p[[q]].$$

**Duality Theorem:** [EMI, §3.2.5]. We have a perfect duality $\langle \cdot, \cdot \rangle : h_k(Np^\alpha; \mathbb{Z}_p) \times H^0(X_1(Np^\alpha)/\mathbb{Z}_p, \omega^k_{\text{cus}}) \to \mathbb{Z}_p$ given by $\langle h, f \rangle = a(1, f|h)$ for $\alpha = 1, \ldots, \infty$, where we put $H^0(X_1(Np^\infty)/\mathbb{Z}_p, \omega^k_{\text{cus}}) := V_k$. So, $V_k \cong V_2^{\text{Katz}} V(N) \subset \mathbb{Z}_p[[q]]$ by Independence theorem.
§6. Ordinary part. The algebra \( h \) is a little too big to have an exact structure theorem. We consider the ordinary projection \( e := \lim_{n \to \infty} U(p)^n! \) inside \( h \). Write \( h_{k}^{ord} = eh_k \) and \( h^{ord} = eh \).

Why \( e \) cuts down \( V(N) \) and \( h \) to a reasonable size?

For whatever \( \alpha > 0 \), we have

\[
\Gamma_0(p^\alpha) \left( \begin{array}{cc} 1 & 0 \\ 0 & p \end{array} \right) \Gamma_0(p^\alpha) = \Gamma_0(p^\alpha) \left( \begin{array}{cc} 1 & 0 \\ 0 & p \end{array} \right) \Gamma_0(p^{\alpha-1}) = \bigcup_{u=0}^{p-1} \Gamma_0(p^\alpha) \left( \begin{array}{cc} 1 & u \\ 0 & p \end{array} \right)
\]

independent of \( \alpha \); so, a power of \( U(p) \) reduces the level of \( f \in S_k(\Gamma_1(p^\alpha)) \) down to \( \Gamma_1(p) \) as long as the Neben character is trivial (contraction property of \( U(p) \)).

**Control theorem:**

\[
h^{ord}/(\langle \gamma \rangle - \gamma^n)h^{ord} \cong h_{k}^{ord}(NP; \mathbb{Z}_p) \quad (k = n + 2 \geq 2).
\]

This is, for example, proven as [EMI, Theorem 4.1.29], and hence \( h^{ord} \) is \( \Lambda \)-free of rank equal to \( h_2(Np; \mathbb{Z}_p) \) by ring theory.
§7. Galois representations.
Similarly $TJ_0$ is also $\Lambda$-free of rank equal to \(\text{rank}_{\mathbb{Z}_p} eH_1^1(X_1(Np), \mathbb{Z}_p)\), which is a Galois module (Tate module of the Jacobian of $X_1(Np)$). By Eichler–Shimura, $\text{Frob}_l$ satisfies $X^2 - T(l)X + l\langle l \rangle = 0$ on $TJ_0$ [GME, Theorem 4.2.2]. Indeed, $TJ_0^{\text{ord}} := eTJ_0$ fits into the following connected-étale exact sequence of $\mathfrak{h}$-modules [H13, Lemma 4.2]:

$$0 \to \mathfrak{h}^{\text{ord}} \to TJ_0^{\text{ord}} \xrightarrow{\text{red}} \text{Hom}_{\Lambda}(\mathfrak{h}^{\text{ord}}, \Lambda) \to 0.$$ 

This follows by a property of reduction modulo $p$ of $X_1(Np^\alpha)$. Take a local ring $\mathbb{T}$ of $\mathfrak{h}^{\text{ord}}$. If $\mathbb{T} \cong \text{Hom}_{\Lambda}(\mathbb{T}, \Lambda)$ (i.e., $\mathbb{T}$ is Gorenstein), we have $TJ_0^{\text{ord}} \otimes_{\mathfrak{h}^{\text{ord}}} \mathbb{T} \cong \mathbb{T}^2$ and we get a Galois representation unramified outside $Np$: $\rho_\mathbb{T} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{T})$, and if $\mathbb{T}$ is not Gorenstein, localize $\mathbb{T}$ into $\mathbb{T}_P$ at a prime $P$ of $\mathbb{T}$ to make $\mathbb{T}_P$ Gorenstein, we have $\rho_\mathbb{T} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{T}_P)$ such that

$$\det(1 - \rho_\mathbb{T}(\text{Frob}_l)X) = 1 - T(l)|_{\mathbb{T}}X - l\langle l \rangle X^2$$

for all primes $l \nmid Np$, where $\langle l \rangle$ is the image of $l$ in $\varprojlim_\alpha \Gamma_0(Np^\alpha)/\Gamma_1(Np^\alpha) \cong \mathbb{Z}_p^\times \times (\mathbb{Z}/N\mathbb{Z})^\times$. 
§8. $\Lambda$-adic forms.
Pick $\lambda \in \text{Hom}_\Lambda(h^{\text{ord}}, \Lambda)$. Consider a formal expansion $F_\lambda := \sum_{n=1}^\infty \lambda(T(n))q^n \in \Lambda[[q]]$. By the duality and control theorem,

$$F_\lambda \mod (\langle \gamma \rangle - \gamma^n) = F_\lambda(\gamma^n - 1)$$

$$= \sum_{n=1}^\infty \lambda(T(n))(\gamma^n - 1)q^n$$

$$\in H^0_{\text{cusp}}(X_1(Np)/\mathbb{Z}_p, \omega^k_{\text{cusp}}) =: S_k(\Gamma_0(Np); \mathbb{Z}_p).$$

Thus $F_\lambda$ is a $\Lambda$-adic form, and

$$\text{Hom}_{\Lambda-\text{alg}}(h^{\text{ord}}, \Lambda) \otimes_{\Lambda} \Lambda / (\langle \gamma \rangle - \gamma^n) \cong S_k(\Gamma_0(Np); \mathbb{Z}_p).$$

If the local ring $\mathbb{T}$ in §7 is equal to $\Lambda$ and $\lambda$ is a $\Lambda$-algebra homomorphism, then

$$\rho_\mathbb{T} \mod (\langle \gamma \rangle - \gamma^n) : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{Z}_p[\psi])$$

is the Galois representation of the Hecke eigenform $F_\lambda(\gamma^n - 1)$. 