Notes for $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-reps representations

1 Motivation

In this last note we explained what types of representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ were important to us (the irreducible+smooth/irreducible+admissible) ones. We said as a motivating factor for this that we were interested in studying 'representations' of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ (where, again, we remark that these are not honest representations of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ but $((\mathfrak{gl}_2)_\mathbb{C}, O_2(\mathbb{R})) \times \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-modules) and, in particular, how these (conjecturally) related to 2-dimensional Galois representations of $G_\mathbb{Q}$.

Now, going from representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ straight to $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$-reps is a large leap, especially because the complicating advent of $((\mathfrak{gl}_2)_\mathbb{C}, O_2(\mathbb{R}))$-modules. So, instead of jumping straight from the purely local picture to the global picture we instead take an intermediate route and study representations of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$. These have the advantage that, formally, they are much like the case of representations of $\mathrm{GL}_2(\mathbb{Q}_p)$—this is because $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ is a TD group.

That said, this intermediary step is not one purely of metered convenience—it's not just so we don't get overwhelmed with the full definition of a $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$-representations—there is serious, valuable intuition about the classic theory of modular forms contained in the study of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-representations. Particularly, one can use this representation theoretic perspective to illuminate the definition of Hecke operators (which, upon a first encounter, seem somewhat... unmotivated) and explain what precisely is the intrinsic significance of phrases like 'Hecke eigenform' and 'newform'.

2 Hecke algebras

2.1 Basic definition

One thing that we neglected to discuss in the last note is the notion of Hecke algebras which are an invaluable tool in the study of smooth representations of TD groups. The idea is simple through the following analogy: Hecke algebras are to TD groups as group algebras are to finite groups. Namely, the Hecke algebra $\mathscr{H}(G)$ of a TD group G is made so that, essentially $\mathsf{Rep}^{\mathrm{sm}}(G) = \mathsf{Mod}^{\mathrm{sm}}(\mathscr{H}(G))$ where the superscript sm means 'smooth' (we'll explain what a smooth $\mathscr{H}(G)$-module is shortly).

So, without further adieu, let G be a TD group. We can then glibly define $\mathscr{H}(G)$ to be the convolution algebra of bi-invariant (under some compact open $K \subseteq G$) measures on G. But, once we fix a Haar measure μ on G (which we do!) we can write it down in more down-to-earth terms. Namely, we define the *Hecke algebra* of G, denoted $\mathscr{H}(G)$, to be the $\mathbb{C}$-space $C_c^\infty(G)$ of compactly supported locally constant functions $f : G \to \mathbb{C}$ with the obvious addition and multiplication given by convolution:

$$(f_1 * f_2)(g) = \int_G f_1(x) f_2(x^{-1}g) d\mu$$

for $f_1, f_2 \in C_c^\infty(G)$.

Remark 2.1: The identification of $\mathscr{H}(G)$ with bi-invariant measures on G follows, essentially, from the Radon-Nikodym theorem—although there is probably a simpler proof here. ◆

We note that if G is not compact then $\mathscr{H}(G)$ is not necessarily unital. Indeed, if G is compact then a unit is certainly given by $\dfrac{1}{\mu(G)}\mathbb{1}_G$. Now while this element obviously won't work for general G (one reason, amongst many, is that $\mathbb{1}_G$ is then not compactly supported) it's maybe slightly non-obvious whether $\mathscr{H}(G)$

is unital. That said, one can confirm one's suspicion—$\mathscr{H}(G)$ is indeed unital if and only if G is compact. This non-unitalness in general makes module theory over $\mathscr{H}(G)$ slightly more nuanced.

Now, if K is a compact open subset of G we denote by $\mathscr{H}(G,K) \subseteq \mathscr{H}(G)$ the subalgebra of those compactly supported locally constant functions $f : G \to C$ which are bi-invariant under K (i.e. invariant under right or left multiplication by K). Note that $\mathscr{H}(G,K)$ *is* a unital algebra with unit $\dfrac{1}{\mu(K)}\mathbb{1}_K$. Indeed, if $f \in \mathscr{H}(G,K)$ then

$$\begin{aligned}\left(\frac{1}{\mu(K)}\mathbb{1}_K * f\right)(g) &= \int_G \frac{1}{\mu(K)}\mathbb{1}_K(x)f(x^{-1}g)d\mu\\ &= \frac{1}{\mu(K)}\int_K f(x^{-1}g)d\mu\\ &= \frac{1}{\mu(K)}\int_K f(g)d\mu\\ &= f(g)\end{aligned}$$

and similarly for right multiplication by $\dfrac{1}{\mu(K)}\mathbb{1}_K$.

Now what makes the study of $\mathscr{H}(G)$-modules reasonable is that while it is not unital (in general) it is a colimit of unital things. Namely, we claim the following:

Theorem 2.2: *Let G be a TD group. Then*

$$\mathscr{H}(G) = \varinjlim \mathscr{H}(G,K)$$

where, of course, if $K \subseteq K'$ then $\mathscr{H}(G,K') \subseteq \mathscr{H}(G,K)$.

Proof: Indeed, this requires just showing that if $f : G \to \mathbb{C}$ is compactly supported and locally constant then f is bi-invariant by some compact open of G. To see this, note that for each point $x \in \mathrm{Supp}(f)$ w we can find a neighborhood of f where f is constant. But, a neighborhood basis of x is (since G is TD) given by right translations by compact open subgroups. Thus, for each x there is some compact open K_x such that $f\mid_{xK_x}$. Since $\mathrm{Supp}(f)$ is compact we can cover it by finitely many xK_x's and thus, evidently, the intersection over this set of finitely many K_x is a compact open subgroup for which f is right invariant. Do the same thing for left invariance and intersect them—this shows the claim! ■

In fact, we can even describe $\mathscr{H}(G,K) \subseteq \mathscr{H}(G)$ in a purely algebraic fashion. Namely, for notational convenience, let us set $e_K := \dfrac{1}{\mu(K)}\mathbb{1}_K$ for every compact open $K \subseteq G$. We then have the following simple proposition that we leave to the reader:

Theorem 2.3: *Let $f \in \mathscr{H}(G)$. Then, f is right K-invariant if and only if $f * e_K = f$ and f is left K-invariant if and only if $e_K * f = f$.*

From this we deduce that $\mathscr{H}(G,K)$ is *precisely* $e_K\mathscr{H}(G)e_K$. Thus, $\mathscr{H}(G)$ is a (possibly) non-unital algebra with a family of idempotents $\{e_K\}$ such that

$$\mathscr{H}(G) = \bigcup_K e_K\mathscr{H}(G)e_K$$

making $\mathscr{H}(G)$ a so-called *idempotented algebra*. The algebra of idempotented algebras, in particular their module theory, is largely translatable from the classical theory.

2.2 Relationship to representation theory

Now that we have a bare-bones understanding of Hecke algebras we can explain why they're of interest to us—how they are related to representation theory. We intuited at the beginning of this section that they should play a role analogous to that played by the group algebra in the case of finite groups.

One reason that this should not be surprising in the slightest is the following:

Example 2.4: Let G be a finite group (considered as a TD group with the discrete topology). Then, $\mathscr{H}(G) = \mathbb{C}[G]$. ▌

But, beyond this literal connection we can actually justify the claim $\mathsf{Rep}^{\mathrm{sm}}(G) = \mathsf{Mod}^{\mathrm{sm}}(\mathscr{H}(G))$ in general. In particular, let us call a $\mathscr{H}(G)$-module M *smooth* if $\mathscr{H}(G)M = \mathscr{H}(G)$. Note that while this is immediate in the module theory of unital algebras (since $1M = M$!) this is not at all a vacuous condition here. That said, one of the key points of noticing that $\mathscr{H}(G)$ is a idempotented algebra is to conclude that, in fact, $\mathscr{H}(G)$ itself is smooth. Namely, if $f \in \mathscr{H}(G)$ then $f \in \mathscr{H}(G, K)$ for some K and thus $e_K * f = f$ so that $f \in \mathscr{H}(G)\mathscr{H}(G)$.

Thus, with this setup we can state the desired result as follows:

Theorem 2.5: *Let G be a TD group. Then, there is a natural equivalence of categories $\mathsf{Rep}^{\mathrm{sm}}(G) \cong \mathsf{Mod}^{\mathrm{sm}}(\mathscr{H}(G))$.*

Here, of course, $\mathsf{Rep}^{\mathrm{sm}}(G)$ denotes the category of smooth G-representations of G (a la the last note) and $\mathsf{Mod}^{\mathrm{sm}}(\mathscr{H}(G))$ denotes the category of smooth $\mathscr{H}(G)$-modules.

Of course, it's not overwhelmingly useful to know this equivalence in the abstract, and so we'd like to make more explicit precisely what the equivalence is. So, let's suppose that V is a smooth G-representation. We then define the structure of $\mathscr{H}(G)$-module on V, described as a homomorphism $\pi : \mathscr{H}(G) \to \mathrm{End}(V)$, as follows:

$$\pi(f)(v) := \int_G f(g)g(v)d\mu \tag{1}$$

While this seems a bit scary looking it's really quite tame. Namely, since V was assumed smooth we know that $\mathrm{stab}(v) \subseteq G$ is open and so contains a compact open subgroup K_1. Moreover, since $f \in \mathscr{H}(G)$ we know that f is bi-invariant for some compact open subgroup K_2 of G. Let $K := K_1 \cap K_2$. Then, one can check that

$$\pi(f)(v) = \mu(K)\sum_i f(g_i)g_i(v) \tag{2}$$

if g_iK is a disjoint open cover of $\mathrm{Supp}(f)$.

Conversely, suppose that V is a smooth $\mathscr{H}(G)$-module (again presented as a homomorphism $\pi : \mathscr{H}(G) \to \mathrm{End}(V)$) we then want to produce a smooth G-module V. Set, not shockingly, V (as an underlying $\mathbb{C}$-space) to be just itself and let us define the G-module structure on V as follows. So, if $v \in V$ then the fact that V is a smooth $\mathscr{H}(G)$-module implies that $v = \pi(f)(w)$ for some $w \in V$ and $f \in \mathscr{H}(G)$. Suppose that $K \subseteq G$ is compact open such that $f \in \mathscr{H}(G, K)$. Then we define $gv := \pi\left(\frac{1}{\mu(K)}\mathbb{1}_{KgK}\right)(v)$ which one can quickly check is, indeed, well-defined.

Let us just quickly verify that these operations are inverse to one another. Namely, suppose that V is a G-module. Then, we want to verify that $\pi(\mathbb{1}_{KgK})(v) = g(v)$. That said, by (2) we have that

$$\pi\left(\frac{1}{\mu(K)}\mathbb{1}_{KgK}\right))(v) = \mu(K)\frac{1}{\mu(K)}\mathbb{1}_{KgK}(g)g(v) = g(v)$$

as desired. Checking the other side of the inverse is exactly the same.

From this more explicit description of the equivalence we can easily suss out the following:

Theorem 2.6: *Let G be a TD group, V a smooth G-rep, and $K \subseteq G$ a compact open subgroup. Then, the following holds:*

1. *The operator $\pi(e_K)$ is the projection operator $V \to V^K$.*
2. *The representation V is admissible if and only if for all $f \in \mathscr{H}(G)$ the endomorphism $\pi(f)$ of V has finite rank.*
3. *The G-stable subspaces of V are the $\mathscr{H}(G)$-submodules of V.*
4. *The $\mathbb{C}$-subspace $V^K \subseteq V$ is a $\mathscr{H}(G, K)$-module.*

Essentially what we see from this theorem is that since $V = \bigcup_K V^K$, V^K is a $\mathscr{H}(G,K)$-module, and $\mathscr{H}(G) = \varinjlim \mathscr{H}(G,K)$ that a smooth G-module V is nothing more than a 'compatible family' of $\mathscr{H}(G,K)$-modules, and that admissibility is that each of these $\mathscr{H}(G,K)$-modules are finite dimensional (over $\mathbb{C}$). In particular, the fact that $\mathscr{H}(G) = \varinjlim \mathscr{H}(G,K)$ is essentially the equality $V = \bigcup_K V^K$ with $V = \mathscr{H}(G)$ acting on itself (the regular representation).

Smooth $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^\infty)$-reps and a baby Flath's theorem

We would now like to begin our study, in earnest, of the admissible irreducible representations of the TD group $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^\infty)$. The key observation is that such representations are 'built from' local pieces. To make this precise we need to make a brief detour to discuss restricted tensor products. Of course, as mentioned in the last note, the fact that $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^\infty) = \prod_p{}' \mathrm{GL}_2(\mathbb{Q}_p)$ should mean that its smooth irreducible representations should decompose as $V = \bigotimes_p{}' V_p$ with V_p a smooth irreducible representation of $\mathrm{GL}_2(\mathbb{Q}_p)$. Of course, step one to understanding why this is true is to understand what $\bigotimes_p{}'$ even means.

So, let us suppose that S is some set of indices and suppose that for every $s \in S$ we have a $\mathbb{C}$-space V_s. Let us define for a finite subset $T \subseteq S$ the space $V_T := \bigotimes_{t \in T} V_t$. What we essentially want to define $\bigotimes_s{}' V_s$ to be is $\varinjlim_T V_T$ as T runs over the finite subsets of S. Of course, for this work then we need for $T' \supseteq T$ a map $V_T \to V_{T'}$. So, let us suppose that for all $s \in S$ we've fixed vectors $0 \neq v_s^0 \in V_s$. Then, let's define the map $V_T \to V_{T'}$ to be $v \mapsto v \otimes v_{s_1}^0 \otimes \cdots \otimes v_{s_m}^0$ if $\{s_1, \ldots, s_m\} = T' - T$. We then set, rigorously, $\bigotimes_s{}' V_s$ to be $\varinjlim_T V_T$. We call it the *restricted tensor product* of the family $\{V_s\}_{s \in S}$.

Remark 2.7: The above can be made slightly nicer by choosing an ordering on T for, as of now, we've not specified the order of the tensor products in V_T. This is not a detail we'll concern ourselves with. ⧫

Since this will also be useful for us let us consider how we might define restricted tensor products of *algebras*. Namely, let us now assume that for each $s \in S$ we're given instead of just a $\mathbb{C}$-space V_s a $\mathbb{C}$-algebra A_s. We then want to define $\bigotimes_s{}' A_s$ in the exact same way but, of course, we'd like to get an algebra out of it. The key is that we can't arbitrarily choose $v_s^0 \in A_s$ else the maps $A_T \to A_{T'}$ won't be multiplicative. What we need is that each v_s^0 is actually an *idempotent* (not zero of course). So, let's suppose that we are given idempotents $e_s^0 \in A_s$ for all s. Then, we can define the *restricted tensor product algebra* $\bigotimes_s{}' A_s$ as $\varinjlim_T A_T$ which is, indeed, a $\mathbb{C}$-algebra since each $A_T \to A_{T'}$ is a map of $\mathbb{C}$-algebras.

Before we give an example, let us point out the obvious—what the *universal property* of this restricted tensor product algebra/vector space is. A $\mathbb{C}$-algebra homomorphism $\bigotimes_s{}' A_s \to B$ (as written on the tin of being a tensor product/colimit) is, for every finite T, a set of maps $f_t : A_t \to B$, which is compatible as T gets larger. In essence, the restricted product of algebras is something like an infinite coproduct (where a map from it is just a collection of maps) but even though all the algebras A_s are taken into account, one only sees finitely many algebras at a given time—one cannot specify the images of $a_s \in A_s$ for *infinitely many* s.

So, what is an example of an algebra that is like 'putting together a bunch of algebras, but only finitely many variables are involved at a given point':

Example 2.8: The polynomial algebra $\mathbb{C}[T_1, T_2, \ldots]$ in infinitely many variables is $\bigotimes_n{}' \mathbb{C}[T_n]$ where the idempotents we take are, not shockingly, the identity element. In general the restricted tensor product of

$A_n := \mathbb{C}[T_{i,n}]/(f_{i,j})$ should be something like $\mathbb{C}[\{T_{i,n}\}_{i,n}]/(\{f_{i,j}\}_{i,n})$. ▮

Now we'd like to understand how one can use restricted tensor products to understand representations of TD groups obtained as restricted direct products. The connection is somewhat clear. Namely, let's suppose that G is a TD group and we can write $G = \prod_s{}' G_s$ with TD groups G_s with respect to some compact open subgroup $H_s \subseteq G_s$. Note then that if we start with representations V_s of each G_s we'd like to say that $\bigotimes_s{}' V_s$ is a representation of G_s. Now, of course, for each finite $T \subseteq S$ we have that V_T is a representation of $G_T := \prod_{t \in T} G_T \times \prod_{t \notin T} H_t$ and since $G = \varinjlim_T G_T$ we seem golden. Of course, this is only if under the transition maps $G_T \to G_{T'}$ we have that the associated maps $V_T \to V_{T'}$ are intertwining. A little thought shows that this is true precisely when the vectors v_0^s we've chosen are H_s stable for all S. In this case we see that $\prod_s{}' V_s$ is a G-rep.

Note that since we need only need to have $H_s \subseteq G_s$ given for all but finitely many s one sees that the discussion of the previous discussion extends to the case when we ignore finitely many $s \in S$, and thus we can ignore picking $v_s^0 \in V_s^{H_s}$ for finitely many s.

We now come to statement of Flath's theorem:

Theorem 2.9 (Baby Flath): *Let V be an irreducible admissible representation of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^\infty)$. Then, there exists a unique decomposition $V \cong \bigotimes_p{}' V_p$ where V_p is an admissible irreducible representation of $\mathrm{GL}_2(\mathbb{Q}_p)$ and V_p is unramified for almost all p (recall that this means that $V_p^{\mathrm{GL}_2(\mathbb{Z}_p)} \neq 0$).*

One should note that, while we have been sloppy suppressing the dependence on the vectors $\{v_s^0\}$ in the notation $\bigoplus_s{}' V_s$, it is now actually important. Namely, Theorem 2.9 is stated in such a way that it seems to *not* depend on a chocie of vectors (up to isomorphism) and this is, indeed, the case.

That said, the justification of this result requires a bit of work. The key is that we must take our vectors $v_0^p \in V_p^{\mathrm{GL}_2(\mathbb{Z}_p)}$, and we do so only when $V_p^{\mathrm{GL}_2(\mathbb{Z}_p)} \neq 0$ (which by the statement of Flath's theorem is non-zero for all but finitely many p). Thus, any ambiguity in the statement of Flath's theorem comes from the non-specification of how we choose these v_0^p and whether this restricted tensor product depends on this choice.

Thankfully, it does not depend on such a choice. The key is the following:

Theorem 2.10 (Baby Satake isomorphism): *Let $T_p := \mathbb{1}_{K_1}$ and $S_p := \mathbb{1}_{K_2}$ where*

$$K_1 = \mathrm{GL}_2(\mathbb{Z}_p)\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}\mathrm{GL}_2(\mathbb{Z}_p)$$

and

$$K_2 = \mathrm{GL}_2(\mathbb{Z}_p)\begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}\mathrm{GL}_2(\mathbb{Z}_p)$$

Then

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), \mathrm{GL}_2(\mathbb{Z}_p)) = \mathbb{C}[T_p, S_p, S_p^{-1}]$$

In particular, we see that $\mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), \mathrm{GL}_2(\mathbb{Z}_p))$ is commutative. Now, if $V_p^{\mathrm{GL}_2(\mathbb{Z}_p)} \neq 0$ then, since V_p is irreducible, this $\mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), \mathrm{GL}_2(\mathbb{Z}_p))$-module (by Schur's lemma) must be 1-dimensional. Thus any two choices of v_p^0 differ by a constant, and one can then show that the isomorphism class of the restricted tensor product is unaffected by this choice.

We will not prove Flath's theorem here referring the interested reader to The Corvallis Proceedings and Flath's orginal article. The result, while deep, is fairly easy to prove relying mostly on basic algebra.

Remark 2.11: Note that the irreducibility of V in Flath's theorem is *pivotal*. Namely, one cannot attempt to decompose a general representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ into local factors. In particular, one cannot apply this to the space $\mathscr{A}_0(\mathrm{GL}_2)$ of GL_2-automorphic forms to obtain 'local automorphic forms'. This is somewhat deep problem, to find what a 'local analogue' of automorphic forms should be. ♦

Let us end this section by stating what is, morally, an equivalent statement of Theorem 2.9 but on the level of Hecke algebras (but it is easier—it's actually a first step into proving Flath's theorem):

Theorem 2.12: *There is an isomorphism of $\mathbb{C}$-algebras*

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)) \cong {\bigotimes_p}' \mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p))$$

here the idempotent we take each $\mathscr{H}(\mathrm{GL}_2(\mathbb{Z}_p))$ is $\mathbb{1}_{\mathrm{GL}_2(\mathbb{Z}_p)}$ (where, implicitly, we've normalized the Haar measure on $\mathrm{GL}_2(\mathbb{Q}_p)$ such that $\mathrm{GL}_2(\mathbb{Z}_p)$ has measure 1).

Moreover, suppose that $U \subseteq \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ is a compact open with factorization $U = \prod_p U_p$. Then

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U) \cong \bigotimes_p \mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), U_p)$$

3 Connection to modular forms

3.1 Basic setup

Let us now give a non-trivial examples of smooth $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-representations and, along the way, clarify the basic theory of modular forms.

The first step is to generalize the notion of a modular form of weight k and level Γ (for Γ a congruence subgroup) to work with level *any* compact open subgroup $U \subseteq \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$. The reason why this should seem like a generalization (and we'll see later is literally a generalization) is that the congruence subgroups $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ are precisely subgroups of $\mathrm{SL}_2(\mathbb{Z})$ of the form $U \cap \mathrm{SL}_2(\mathbb{Z})$ with $U \subseteq \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ compact open.

So, without further delay let us define such a modular form. Namely, let us say that a function

$$f : \mathfrak{h}^\pm \times \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty) \to \mathbb{C}$$

(where $\mathfrak{h}^\pm$ is the upper and lower half-planes) is a *modular form of weight k and level U*, where $U \subseteq \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ is compact open, if it satisfies the following conditions:

MF.1 For each fixed $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ we have that $f(h, g)$ is a holomorphic function $\mathfrak{h}^\pm \to \mathbb{C}$.

MF.2 For each $\gamma \in \mathrm{GL}_2(\mathbb{Q})$ we have that

$$f(\gamma h, \gamma g) = \deg(\gamma)^{-1} j(\gamma, z)^k f(h, g)$$

where $\mathrm{GL}_2(\mathbb{Q})$ acts on $\mathfrak{h}^\pm$ by fractional linear transformation and on $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ diagonally, and where $j\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, z\right) = cz + d$.

MF.3 For all $u \in U$ we have that $f(h, gu) = f(h, g)$.

MF.4 For all fixed $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ we have that $\lim_{h \to i\infty} f(h, g)$ exists.

Condition **MF.4** can be rephrased in two useful ways. First, it is equivalent to say that (again for each fixed g) for every $A \geqslant 0$ there is a $C_A > 0$ and $N \in \mathbb{N}$ such that $|f(h, g)| \leqslant C_A \operatorname{Im}(h)^N$ when $|\mathrm{Re}(x)| \leqslant A$ and $\operatorname{Im}(h) \gg 0$—in other words, that $f(g, h)$ is of *moderate growth*. Second, note that $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in U$ for $N \gg 0$ which shows that $f(h, g) = f(h + N, g)$ so that f still has a Fourier expansion

$$f(h, g) = \sum_{n \in \mathbb{Z}} a_n \exp\left(\frac{2\pi i h}{N}\right) \tag{3}$$

and we require that $a_n = 0$ for $n < 0$.

If, in addition, we assume that for each g we have that $\lim_{h\to i\infty} f(h,g) = 0$ then we say that f is a *cuspform of weight k and level U*. Equivalently, we can say that, for each fixed g, the function $f(h,g)$ is of *rapid decay* meaning that for each fixed A and $k \in \mathbb{N}$, there is a $C_{A,k}$ such that $|f(h,g)| \leqslant C_A \operatorname{Im}(h)^{-k}$ for $|x| \leqslant A$ and $\operatorname{Im}(h) \gg 0$. Finally, we can require that for each fixed g the coefficient a_0 in (3) is 0.

One can think about what we've done above as follows. It's a common first exercise in the study of automorphic forms to explain how one can inflate functions $f : \mathfrak{h} \to \mathbb{C}$ to functions $f : \mathrm{SL}_2(\mathbb{R}) \to \mathbb{C}$ which begins one on the journey of understanding 'automorphic forms for $\mathrm{SL}_2(\mathbb{R})$' and how they generalize classic modular forms. Here we have, instead, inflated our modular forms to live on *non-archimedean groups* (or rather, those that have a large non-archimedean component) and left the archimedean factor alone. This is why, for instance, prescriptions of how $\mathrm{O}_2(\mathbb{R})$ acts are not necessary—we've done no inflating on the Archimdean factor.

One of the powers of this approach is that it allows us to unite the study of modular forms and cusp forms across different weights together. Namely, let us denote for each compact open U the $\mathbb{C}$-space $M_k(U)$ of modular forms of weight k and level U and similarly we denote cuspforms by $S_k(U)$. Note that if $U' \supseteq U$ then evidently $M_k(U') \subseteq M_k(U)$ and $S_k(U') \subseteq S_k(U)$. So, let us define the following:

$$\mathcal{M}_k := \varinjlim_U M_k(U)$$

$$\mathcal{S}_k := \varinjlim_U S_k(U)$$

And note that $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ now actually *acts* on these spaces. Specifically, if $g_0 \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ and we define, as per usual, the function $g_0 f$ by the rule

$$(g_0 f)(h,g) = f(h, gg_0)$$

then for any $f \in M_k(U)$ we have that $g_0 f \in M_k(g_0^{-1}Ug_0)$ and if $f \in S_k(U)$ then $g_0 f \in S_k(g_0^{-1}Ug_0)$. Thus, while we can't see something like a $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-action at a fixed level (we'll see shortly what type of action we really get there) once we unify the weight k forms across all levels the $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ action becomes apparent.

Before we continue, let us prove a small lemma that will show in many cases that $M_k(U)$ and $S_k(U)$ are familiar objects:

Theorem 3.1: *Let $U \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}})$ be a compact open. Suppose moreover that $\det(U) = \widehat{\mathbb{Z}}^\times$. Then, if $\Gamma := \mathrm{SL}_2(\mathbb{Z}) \cap U$ then Γ is a congruence subgroup and the map $f \mapsto f(h,1)$ defines a bijection $M_k(U) \xrightarrow{\approx} M_k(\Gamma)$ and $S_k(U) \xrightarrow{\approx} S_k(\Gamma)$.*

This theorem is fairly simple, but let us remark on the inclusion of the condition $\det(U) = \widehat{\mathbb{Z}}^\times$. Recall that the algebraic group GL_2 does not satisfy strong approxmation. Thus, to prove such a theorem we need to bootstrap from strong approximation for SL_2 and the condition that $\det(U) = \widehat{\mathbb{Z}}^\times$ essentially allows us to make the equality

$$\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)/U = \mathrm{SL}_2(\mathbb{A}_\mathbb{Q}^\infty)/(U \cap \mathrm{SL}_2(\mathbb{A}_\mathbb{Q}^\infty))$$

which is what makes this possible.

Remark 3.2: Of course, this then wonders why we're not working with just SL_2 instead of GL_2 where the above theorem wouldn't need such qualifications. There are two answers. First is, well, we're interested in $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ (because it's on the seminar organizer's qualifying exam). Secondly, it's GL_n that much of what we say extends to. Namely, we'll use below heavily the fact that GL_2 has the strong multiplicity one property. This also holds for GL_n. That said, while it holds for SL_2 it does *not* hold for SL_n for $n > 2$. ♦

In particular, let's apply Theorem 3.1 to some familiar situations. Namely, let us consider the following open subgroups of $\mathrm{GL}_2(\widehat{\mathbb{Z}})$:

$$U(N) := \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) : a \equiv 1 \mod N,\, b \equiv c \equiv 0 \mod N\right\}$$
$$U_1(N) := \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) : a \equiv 1 \mod N,\, c \equiv 0 \mod N\right\}$$
$$U_0(N) := \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) : c \equiv 0 \mod N\right\}$$

Then these all satisfy $\det(U) = \widehat{\mathbb{Z}}^\times$. Moreover, one can check that

$$U(N) \cap \mathrm{SL}_2(\mathbb{Z}) = \Gamma(N)$$
$$U_1(N) \cap \mathrm{SL}_2(\mathbb{Z}) = \Gamma_1(N)$$
$$U_0(N) \cap \mathrm{SL}_2(\mathbb{Z}) = \Gamma_0(N)$$

and thus we have equalities

$$M_k(U(N)) = M_k(\Gamma(N)) \qquad S_k(U(N)) = S_k(\Gamma(N))$$
$$M_k(U_1(N)) = M_k(\Gamma_1(N)) \qquad S_k(U_1(N)) = S_k(\Gamma_1(N))$$
$$M_k(U_0(N)) = M_k(\Gamma_0(N)) \qquad S_k(U_0(N)) = S_k(\Gamma_0(N))$$

thus connecting this theory back to the most basic cases of modular forms.

Remark 3.3: Be careful! Note that a more natural choice for compact open subgroups of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ giving the $\Gamma_i(N)$ would have been $\widehat{\Gamma_i(N)}$—their profinite completions. These are certainly perfectly fine compact open subgroups of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ but they do *not* satisfy the surjectivity condition for the determinant, thus one cannot directly relate modular forms of their level to modular forms (in the classical sense) of level $\Gamma_i(N)$. In particular, since the canonical neighborhood basis of the identity is $\widehat{\Gamma(N)}$ note that we cannot, *a priori*, uniformize all 'generalized modular forms' (with levels compact opens of $\mathrm{GL}(\mathbb{A}_\mathbb{Q}^\infty)$) in terms of classical modular forms.

Namely, it's not true that we can describe $\mathcal{M}_k$ and $\mathcal{M}_K$ as a colimit of classical modular form spaces. This should be clear since such a space would not have a $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ action (just a $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ action!). So, we have also really gained something by passing to this more general level. ♦

So let us begin by making the following two observations which, while not overly deep, are also not extremely obvious:

Theorem 3.4: *For all $U \subseteq \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ compact open the $\mathbb{C}$-space $M_k(U)$, and thus $S_k(U)$, are finite-dimensional.*

Proof (idea): Just like in the classical case one can write down a (possibly disconnected) compact Riemann surface for which $M_k(U)$ and $S_k(U)$ are sections of a line bundle, thus finite-dimensional. ■

We also need the following:

Theorem 3.5: *There is a natural inner product on $\mathcal{S}_k$ by which $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ acts unitarily (it's, up to a character twist, something like the Peterson inner product).*

So, using Theorem 3.4 we can actual say something somewhat surprising/deep. Namely, the representations $\mathcal{M}_k$ and $\mathcal{S}_k$ of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ are admissible! Indeed, the fact that they are smooth is clear. If $f \in M_k(U) \subseteq \mathcal{M}_k$ then f is fixed by U and thus $\mathrm{stab}(f)$ is open (the same goes for $\mathcal{S}_k$). And, to see it's admissible we note that for each compact open U we have that

$$\mathcal{M}_k^U = M_k(U), \qquad \mathcal{S}_k^U = S_k(U)$$

which are finite-dimensional by Theorem 3.4.

Now that we know that $\mathcal{M}_k$ and $\mathcal{S}_k$ are reasonable representations of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ (they are admissible) we can hope to try and understand their admissible irreducible subrepresentations/subquotients. To this end, let us make the following definition. A *modular representations of weight k* (for GL_2) is an irreducible admissible subquotient of $\mathcal{M}_k$. A *cuspidal representation of weight k* (for GL_2) is an admissible subrepresentation of $\mathcal{S}_k$.

Remark 3.6: Note that in our definition of cuspidal representation of weight k we said subrepresentation instead of subquotient. One can show that the two notions are equivalent. ♦

Thus, the rest of this note will be devoted to understanding precisely the cuspidal representations of weight k.

3.2 Hecke operators

Before we do this directly, it's helpful (and extremely illuminating!) to first understand how the Hecke operators in the classical setting relate to the Hecke algebra in this setting of smooth $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-representations.

So, we said earlier that the individual levels $M_k(U)$ and $S_k(U)$ did not possess a $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-action. So, then, what action do they have? This is a little cumbersome to state in the language of representations but if we think of $\mathcal{M}_k$ and $\mathcal{S}_k$ as smooth $\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty))$-modules the picture suddenly becomes much more sunny. Namely, $M_k(U) = \mathcal{M}_k^U$ is precisely a $\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U)$-module and similarly for $S_k(U)$. Thus, the action that actually exists at a finite level is just the action of the U bi-invariant Hecke algebra $\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U)$.

So, in particular, let us think about the case of

$$M_k(\Gamma_1(N)) = M_k(U_1(n)) = \mathcal{M}_k^{U_1(N)}$$

and similarly for $S_k(\Gamma_1(N))$. Namely, by the above analysis this should carry an action of $\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U_1(N))$. So, for starters, what precisely does this algebra look like? Well, using Theorem 2.12 and Theorem 2.10 we can give a pretty satisfactory answer to this question.

Namely, let us begin by noting that we can write

$$U_1(N) = \prod_{p\nmid N} \mathrm{GL}_2(\mathbb{Z}_p) \times \prod_{p\mid N} U_{1,p}(N)$$

where, here,

$$U_{1,p}(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}_p) : c \equiv 0 \mod p^{v_p(N)},\ a \equiv 1 \mod p^{v_p(N)} \right\}$$

Then, from Theorem 2.12 we deduce that

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U_1(N)) = {\bigotimes_{p\nmid N}}' \mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), \mathrm{GL}_2(\mathbb{Z}_p)) \otimes \bigotimes_{p\mid N} \mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), U_{1,p}(N))$$

That said, we know from Theorem 2.10 that

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), \mathrm{GL}_2(\mathbb{Z}_p)) = \mathbb{C}\left[T_p, S_p, S_p^{-1}\right]$$

and so combining this with Example 2.8 we deduce that

$$\mathscr{H}(\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty), U_1(n)) = \mathbb{C}[\{T_p, S_p, S_p^{-1}\}_{p\nmid N}] \otimes \bigotimes_{p\mid N} \mathscr{H}(\mathrm{GL}_2(\mathbb{Q}_p), U_{1,p}(N))$$

Thus, we see that for each $p \nmid N$ we get operators T_p and S_p on $S_k(\Gamma_1(N)) = \mathcal{S}_k^{U_1(N)}$. The thing we'd hope is, in fact, true:

Theorem 3.7: *Under the identification $S_k(\Gamma_1(N)) = \mathcal{S}_k^{\Gamma_1(N)}$ the operator T_p acts as $p^{1-\frac{k}{2}} T_p^c$ where T_p^c the classical Hecke operator and S_p acts as multiplication by $\chi(p)$ if χ is the Nebentypus for f.*

Of course, one can extend the above result to the Hecke operators at primes $p \mid N$ but that requires more work (in particular, Baby Satake [or Papa Satake] don't apply).

Let's spell out the philosophical implication of all of this. When one is first introduced to Hecke operators, unless they had a particularly good teacher they probably seemed unmotivated and forced. One usually tries to justify them as 'averaging operators' used to prove things about Fourier coefficients. But, the above tells us that if we pay attention to the $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$-action on modular forms (the thing that we care about from a representation theoretic perspective) then the Hecke operators are *forced* on us as being the real 'important part' of this $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)$ action. They don't seem random anymore.

Remark 3.8: One can also understand the Hecke operators naturally from a geometric point of view. Namely, the singular cohomology of any symmetric space has Hecke operators coming from the natural Hecke correspondences of the family of spaces obtained by quotienting this symmetric spaces by congruence subgroups. This is then connected to modular forms via the Matsushima formula and then one obtains, again, the classical Hecke operators. ⧫

3.3 Decomposition $\mathcal{S}_k$

So, in this section we'd like to understand how precisely representations of $\mathcal{S}_k$ look—what cuspidal representations of weight k look like.

The first key observations if the following:

Theorem 3.9: *We have an (algebraic) decomposition*

$$\mathcal{S}_k = \bigoplus_{\substack{V \text{ cusp.} \\ \text{rep. wght } k}} V^{n_V} \tag{4}$$

This is essentially due to Theorem 3.5. Also, note that since $\mathcal{S}_k$ is admissible we know that n_V is finite for all V.

Now, the first big result that we will need (and blackbox heavily) is the following:

Theorem 3.10 (Strong multiplicity 1 for $\mathbf{GL_2}$): *Suppose that V and W are cuspidal representations of weight k embedded into $\mathcal{S}_k$. Then, if $V_p \cong W_p$ for almost all p (where V_p and W_p are as in Theorem 2.9) then $V = W$.*

In other words, for all V we have that $n_V = 1$ in (4) and for V and W occuring in (4) we have that $V = W$ if and only if $V_p \cong W_p$ for almost all p. This is an *extremely* powerful and surprising result and will essentially be the key to all of what we say after.

Remark 3.11: Note that strong multiplicity 1 is a truly global phenomemon. For example, doing the Archimedean analogue of the decomposition in (4) gives arbitrarily high multiplicities. ⧫

So, we now come to what is, in my estimation, one of the most beautiful theorems in the basic theory of automorphic forms/representations which clarifies much of the basic theory. But, before we come to the theorem directly, let us first notate for $f \in S_k(U)$ the orbit $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)f \subseteq \mathcal{S}_k$ by π_f.

We then have the following:

Theorem 3.12: *For $f \in S_k(\Gamma_1(N))$ the subrepresentation $\pi_f \subseteq \mathcal{S}_k$ is irreducible if and only if f is an eigenform. Moreover, for every cuspidal representation V of weight k we have that $V \cong \pi_f$ for some eigenform $f \in S_k(\Gamma_1(N))$ (for some N).*

Proof (Idea): If π_f is irreducible then we can, by Theorem 2.9, decompose it as $\displaystyle\bigotimes_p{}' V_p$ and for almost all p we have that T_p acts by a character, and thus f must be an eigenvalue for T_p.

To see that π_f is irreducible if f is an eigenform we decompose π_f into irreducibles and note that for each irreducible factor its p^{th} Flath component needs to have $p^{\frac{k}{2}-1}\lambda_p(f)$ (where $\lambda_p(f)$ is the eigenvalue of T_p^c on f) as the eigenvalue for T_p. By strong multiplicity 1 this implies that all the irreducible factors are equal and, again, implies that there must only be one irreducible factor.

The fact that all show up as π_f is somewhat obvious. Namley, if V is a cuspidal representation of weight k we know that $V = \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}^\infty)f$ for any non-zero $f \in V$. If we can show that $f \in S_k(\Gamma_1(N))$ then necessarily f is an eigenform since the T_p in the local Hecke algebra must act as a character. Thus, the oomph is showing that we can take $f \in S_k(\Gamma_1(N))$. This we don't do here. ■

Thus, we see that the reason we care about eigenforms is that, well, they are the irreducibles! They are the 'important parts' the 'generating set' of the whole of cuspforms.

We end this section by explaining how newforms (which is short for new normalized Hecke eigencuspform) enter the picture. The basic idea is simple. Namely, we now know that in the decomposition (4) each V is some π_f. Thus, we'd like rewrite this decomposition that form. But, the question is the following: what indexing set of f's should it run over? Namely, there are *many* eigenforms with the same eigenvalues (e.g. $\Delta(z)$ and $\Delta(2z)$ for Δ the usual discriminant form of weight 12 and level $\mathrm{SL}_2(\mathbb{Z})$). The key is that two *newforms* which have the same Hecke eigenvalues are equal.

Thus, the point is that each π_f has a canonical representative given by a newform (and all other Hecke eigenforms are obtained from this on by a $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^{\infty})$-action). Thus, we can rewrite the decomposition (4) as follows:

$$\mathcal{S}_k = \bigoplus_{f \text{ newform}} \pi_f \tag{5}$$

which I find exceedingly beautiful and illuminating, explaining why eigenforms, and in particular newforms, are so important.

3.4 Local factors and L-functions

We end this note by trying to tie the representations π_f, with f a newform for $\Gamma_1(N)$ (for some N), together with the material from the last note. Namely, by Theorem 2.9 we know that

$$\pi_f \cong {\bigotimes_p}' V_{p,f}$$

where V_p is an admissible $\mathrm{GL}_2(\mathbb{Q}_p)$-representation, and so one might begin to wonder what these are.

Well, one case this is clear is in the unramified case. Namely if $p \nmid N$ then it's fairly clear from our analysis above that $V_{p,f}$ must be unramified and thus, from our discussion last time, an irreducible principle series $P(\chi_1, \chi_2)$ and we know that, up to reordering, χ_1 is (in the parlance of last time) the character $(p^{1-\frac{k}{2}} a_p(f), 0, 1)$ and χ_2 the character $(\chi(p), 0, 1)$.

What happens for primes $p \mid N$ is more complicated. Namely, one can show that in this case one never gets a character and thus one is left with either an irreducible (ramified) principal series, a special representation $\mathrm{St}_{\mathrm{GL}_2}$, or a supercuspidal. Loeffler and Weinstein have a paper detailing an algorithm to decide which is which.

Regardless, one thing that one can prove abstractly (without knowing the specific identity of the local representations $V_{f,p}$) is that for every prime p one has that $L_p(f,s) = L_p(\pi_f, s)$ where, by definition, $L_p(\pi_f, s) := L(V_{p,f}, s)$ where we defined last time the definition of the L-function of (non-character) irreducible admissible representation of $\mathrm{GL}_2(\mathbb{Q}_p)$.

Thus, not only do newforms intimately factor into the study of representations of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}}^{\infty})$ but their native L-functions agree with the 'automorphic L-function' obtained by studying their local representations.