THE DEFINITION OF AN AUTOMORPHIC REPRESENTATION (AND HOW TO GET ONE FROM A HOLOMORPHIC FORM)

The references are missing, sorry! Otherwise, please let me know any comments or corrections - AV.

1. Introduction

Given a (connected) reductive group **G** over a number field *F*, and a *coefficient ring R*, an idealized version of the Langlands program posits a bijection between two classes of objects:

- (a) Automorphic forms for G with coefficients in R; this is a generalization of the notion of modular form.
- (b) Generalized Galois representations with targets into $G^{\vee}(R)$, where G^{\vee} is the dual group to **G**.

In general, there are many ways one can go about defining versions of either (a) or (b). We have "satisfactory" definitions for both sides for special types of R, i.e. definitions that subsume all reasonable versions and simultaneously have good formal properties, for example:

- We have a satisfactory notion of "automorphic form" when R = C; it generalizes in a clean way the notion of Hecke eigenform in the theory of modular forms.
 In this document we explain a subcase of it, the notion of *discrete-series* automorphic form.¹
- If R is finite, a generalized Galois representation is just a usual Galois representation, i.e. a homomorphism from the Galois group $Gal(\overline{F}/F)$ to $G^{\vee}(R)$.

But we don't have a clean notion of automorphic form if $R = \mathbf{F}_p$ – at least, not one that would include the somewhat mysterious "even Galois representations" on side(b). Nor do we have a good notion of generalized Galois representation for $R = \mathbf{C}$ – at least not one that would include the "Maass forms" on side (a). Part of the difficulty with writing down exact statements of the Langlands correspondence is that we don't really know a ring R for which both sides behave satisfactorily at the same time: one needs to restrict or enlarge one or both of the sides.

2. Unitary representations

The theory of infinite-dimensional representations of groups has several "avatars." For example, if we consider the action of a Lie group G on the space of functions on G, we have several choices as to how to formalize it. We could use different function spaces (e.g. $L^2(G)$ or $L^p(G)$ or $C^\infty(G)$). We could also try to pass from G to its Lie algebra to get a more algebraic theory. These theories are all very closely linked. For us the shortest route to the definition is to use the *unitary theory*, which we briefly recall. ADDREF

1

¹The adjective *discrete-series* should correspond, under the idealized Langlands program, to the adjective *irreducible* on the other side. Here "irreducible" for a representation with target a general reductive group means: image not contained in a proper parabolic subroup.

Definition. *If* G *is a topological group, then a* unitary representation *of* G *is an isometric action of* G *on a Hilbert space* H *so that the action map* $G \times H \rightarrow H$ *is continuous.*

The basic example is G acting on $L^2(G)$ by left- or right- translation. The notion of continuity above is perhaps not the most obvious one; it's extrapolated from this basic case.

Definition. A unitary representation is said to be irreducible if it admits no nonzero, proper, closed, invariant subspaces.

We will want to use later the following version of Schur's lemma in this world: If G acts irreducibly on the Hilbert space H, and $z: H \to H$ is a *isometry* commuting with G, then z acts by a scalar. This follows as in the usual proof of Schur's lemma, but one replaces the use of eigenspaces by the spectral theorem for unitary operators on a Hilbert space.

In general, a reducible representation need not contain an irreducible; the translation action of \mathbf{R} on $L^2(\mathbf{R})$ gives a counterexample. Indeed, for any measurable set $S \subset \mathbf{R}$, the vector space of functions f for which \widehat{f} has support in S gives an invariant closed subspace. But it has no irreducible subrepresentations: By Schur's lemma (see §2) any irreducible subrepresentation $W \subset L^2(\mathbf{R})$ would have dimension 1, i.e. $W = \langle F \rangle$ where all translates of F are proportional to F. Thus $F \propto e^{\lambda x}$, and this never lies in L^2 .

However a representation "decomposes" into a (possibly uncountable) number of irreducibles: this is the theory of "unitary disintegration," using the notion of direct integral of representations. For \mathbf{R} on $L^2(\mathbf{R})$ this comes down to the theory of Fourier transforms. The existence of unitary disintegration is surprisingly easy; what is not easy is uniqueness [?].

For noncompact groups G, irreducible representations are usually infinite-dimensional and it is often rather difficult to construct any such. We will see some examples later.

3. THE DEFINITION OF AUTOMORPHIC REPRESENTATION

Let F be a global field, with adele ring F. Let G be a reductive group over F. Let $[G] = G(F) \setminus G(A)$; it admit a right action of G(A) and carries a G(A)-invariant measure, unique up to scaling. ADDREF

A basic example to keep in mind is $G = SL_n$ and F = Q; then we will discuss the geometry of $[SL_n]$ in the next section.

Definition. (For **G** semisimple:) A (discrete-series) automorphic representation for **G** is an irreducible unitary subrepresentation of $L^2([\mathbf{G}])$.

We will often drop the "discrete" and refer simply to an automorphic representation, although later the distinction will become important.

Strictly speaking, then, an automorphic representation is a pair:

unitary representation of G(A), isometric embedding $W \to L^2([G])$

up to isomorphism.

For example, the space of *constant* functions is a discrete automorphic representation, because [G] has finite volume for G semisimple $[\ref{G}]$ and so the constant function is square-integrable!

At this stage, it is not at all obvious that there are any *other* discrete automorphic representations; for $F = \mathbf{Q}$ and $\mathbf{G} = \mathrm{SL}_2$ some will be constructed in §5.

3.1. **Comparison with other definitions.** TO BE ADDED (not covered in lecture).

3.2. **The definition for reductive G.** Why the restriction to **G** semisimple? If not, [**G**] may not have finite volume and in that case (by a variant of the argument for **R** acting on $L^2(\mathbf{R})$) has no irreducible subrepresentations at all.²

We can circumvent this problem by working modulo center. Write \mathbf{Z}_G for the center of \mathbf{G} . By Schur's lemma again, an irreducible unitary representation of $\mathbf{G}(\mathbf{A})$ has a central character: i.e. there is a continuous homomorphism $\omega: \mathbf{Z}_G(\mathbf{A}) \to S^1$ with the property that

(1)
$$z \cdot w = \omega(z) w, \ z \in \mathbf{Z}_G(\mathbf{A}), w \in W.$$

Indeed, take any $z \in \mathbf{Z}_G(\mathbf{A})$;

Now fix a character ω as above, and consider $L^2_{\omega}([\mathbf{G}])$: the space of functions f satisfying (1) such that $|f|^2$ is integrable on $[\mathbf{G}]/\mathbf{Z}_G(\mathbf{A})$. A *(discrete) automorphic representation* for \mathbf{G} is an irreducible unitary subrepresentation of $L^2_{\omega}([\mathbf{G}])$. Again, more formally, it would be a pair as above

irreducible unitary representation W, embedding into L_{ω}^2 ,

where ω is the central character of $\mathbf{Z}_G(\mathbf{A})$ on W, characterized by (1).

4. Examples of the automorphic space

We now try to get intuition for the meaning of [G]. As we will see, it is an inverse limit of manifolds of dimension equal to $\dim(G)$, which means in practice that most questions on [G] can be reduced to a question on a finite-dimensional manifold.

For our examples we stick to the case $F = \mathbf{Q}$.

4.1. **[G]** for $G = SL_n$. We claim that there is a homeomorphism

(2)
$$[\mathbf{G}] \simeq \lim \Gamma(N) \backslash \mathrm{SL}_n(\mathbf{R}).$$

Here $\Gamma(N) \subset \operatorname{SL}_n(\mathbf{Z})$ is the principal congruence subgroup, i.e. the kernel of $\operatorname{SL}_n(\mathbf{Z}) \to \operatorname{SL}_n(\mathbf{Z}/n\mathbf{Z})$. These spaces form a projective system, because of the inclusions $\Gamma(M) \subset \Gamma(N)$ whenever M is a multiple of N. We show how to construct the bijection, but leave all checking of topology (continuous, continuous inverse) to the reader.

Let K(N) be the subgroup of $SL_n(\mathbf{A}_f)$ corresponding to $\Gamma(N)$, i.e. the kernel of

$$\mathrm{SL}_n(\widehat{\mathbf{Z}}) \to \mathrm{SL}_n(\mathbf{Z}/N\mathbf{Z}).$$

Here $\widehat{\mathbf{Z}}$ is the profinite completion of \mathbf{Z} , i.e. $\varprojlim \mathbf{Z}/N\mathbf{Z}$; it is identified with the infinite product $\prod_p \mathbf{Z}_p$, and is thus naturally a subring of \mathbf{A}_f – in fact, it is the maximal compact subring. What is important for us is that the K(N) are a cofinal system of compact open subgroups of $\mathrm{SL}_n(\mathbf{A}_f)$.

We claim that the inclusion

$$SL_n(\mathbf{R}) \to SL_n(\mathbf{A}) \to SL_n(\mathbf{Q}) \backslash SL_n(\mathbf{A}) / K(N)$$

actually descends to a bijection $\Gamma(N)\backslash SL_n(\mathbf{R})\simeq SL_n(\mathbf{Q})\backslash SL_n(\mathbf{A})/K(N)$. Injectivity is easy, so just surjectivity: this is the assertion that $SL_n(\mathbf{Q})\cdot SL_n(\mathbf{R})\cdot K(N)$ is all of $SL_n(\mathbf{A})$; or equivalently that $SL_n(\mathbf{Q})\cdot K(N)$ is $SL_n(\mathbf{A}_f)$. But this follows from *strong approximation*: $SL_n(\mathbf{Q})$ is dense in $SL_n(\mathbf{A}_f)$. (Sketch of proof: it suffices to check that the closure of $SL_n(\mathbf{Q})$ contains $SL_n(\mathbf{Q}_p)$ for every finite p; given $g\in SL_n(\mathbf{Q}_p)$, write it as a product of elementary matrices and then use the fact that \mathbf{Q} is dense in \mathbf{A}_f to approximate those. Details of a related argument for *any* split, simply connected group are given in a handout in Brian Conrad's first course on algebraic groups.)

²In fact, [G] fails to have finite volume if and only if there is a central \mathbb{G}_m , see [?].

The statement that $G(Q) \subset G(A_f)$ is dense is *false* for PGL_n ; its validity for SL_n has to do with the fact that SL_n is simply connected.³

In down to earth terms, this says that the inverse limit $\varprojlim \Gamma(N) \backslash SL_n(\mathbf{R})$ inherits an action of a much bigger group than one would expect. On the one hand, $SL_n(\mathbf{Z}/N\mathbf{Z})$ acts on $\Gamma(N) \backslash SL_n(\mathbf{R})$, and so the inverse limit

$$\varprojlim \operatorname{SL}_n(\mathbf{Z}/N\mathbf{Z}) = \operatorname{SL}_n(\widehat{\mathbf{Z}}) \simeq \prod_p \operatorname{SL}_n(\mathbf{Z}_p)$$

acts on the inverse limit. But the claim is that this extends to an action of $SL_n(\mathbf{Q}_p)$. We examine this situation next.

4.2. **The extra action.** We now try to describe what the points of $[SL_n]$ parameterize, and thus the extra action of $SL_n(\mathbf{A}_f)$ just described.

We may bijectively identify

(3)
$$\Gamma(N) \setminus SL_n(\mathbf{R}) \simeq \{L \text{ area 1 lattice in } \mathbf{R}^n, \varphi : L/NL \simeq (\mathbf{Z}/N\mathbf{Z})^n, \cdots \}$$

where \cdots denotes that we need to be careful about determinants: Define a "positive" basis e_1, e_2, \dots, e_n for L to be one such that $\det(e_1, e_2, \dots, e_n) > 0$. Such a basis is unique up to $\mathrm{SL}_n(\mathbf{Z})$; and we require that φ has determinant 1 with respect to e_1, \dots, e_n and the standard basis for $(\mathbf{Z}/N\mathbf{Z})^n$.

Equivalently, we fix an oriented lattice Λ – i.e., a lattice with a basis for $\Lambda^n \Lambda$ chosen, i.e. an equivalence class of "positive bases." Then $\Gamma(N)\backslash \mathrm{SL}_n(\mathbf{R})$ becomes identified with $\mathrm{SL}(\Lambda)$ -equivalence classes of

(4) { oriented isomorphism $\Lambda \otimes \mathbf{R} \simeq \mathbf{R}^n$, oriented isomorphism $\Lambda/N \simeq (\mathbf{Z}/N)^n$ }

where "oriented" means that it has determinant 1 with respect to the chosen class of Λ -bases, and the standard coordinate bases on the right-hand side. To pass from (4) to (3), we take L to be the image of Λ under $\Lambda \otimes \mathbf{R} \simeq \mathbf{R}^n$.

But we can rephrase (5) in a very symmetric way: Start with an n-dimensional **Q**-vector space V, and consider up to SL(V)-equivalence an admissible collection of oriented isomorphisms

(5) { oriented isomorphism
$$\varphi_v : V \otimes \mathbf{Q}_v \simeq \mathbf{Q}_v^n$$
 }

where we include all places v. By *admissible* we mean the following: If $L \subset V$ is any lattice, we require that the closure of $\varphi_v(L)$ be \mathbf{Z}_v^n for all but finitely many v; this requirement doesn't depend on L.

We recover (4) from (5) by taking the lattice Λ inside V by taking the preimage of \mathbf{Z}_p^n for all finite p.

From (5) it is now visible that $SL_n(\mathbf{Q}_p)$ acts on $[SL_n]$ for every p. How do we interpret the action of $g \in SL_n(\mathbf{Q}_p)$? (*To be updated after Oct 22 lecture -AV.*)

4.3. Forcing Hecke operators to be invertible. To paraphrase the last \S slightly, [G] is a version of $G(Z)\setminus G(R)$ where "Hecke operators have become invertible." It's worth noting that one can carry out a general construction of this form:

Given a multivalued function T on a space X, it is always possible to lift X to a space $\tilde{X} \to X$ with a single-valued invertible function $T: \tilde{X} \to \tilde{X}$ that "induces T." Namely, we take \tilde{X} to be the infinite sequence space

$$(\cdots, x_{-2}, x_{-1}, x_0, x_1 x_2, \cdots) : x_{i+1} \in T x_i)$$

 $^{^{3}}$ In fact, it's true exactly when **G** is simply connected and every **Q**-simple factor of **G** is isotropic over **R**. See [?].

and take \tilde{T} to be the shift operator.

If we apply this construction to $X = SL_2(\mathbf{R})/SL_2(\mathbf{Z})$ with all Hecke operators simultaneously, the space \tilde{X} will be closely related to $[SL_2]$.

4.4. **[G]** for $G = GL_n$. Here the situation isn't quite as simple:

$$[GL_n] = \lim[GL_n]/K(N)$$

but each space is actually a union of copies of $\Gamma(N)\backslash \mathrm{GL}_n(\mathbf{R})$ – in fact, $\varphi(N)/2$ such copies. In fact, this space still parameterizes

$$\Lambda \otimes \mathbf{R} \simeq \mathbf{R}^n$$
 together with $\Lambda/N\Lambda \simeq (\mathbf{Z}/N)^n$,

up to $GL(\Lambda)$.

But now $GL(\Lambda)$ acts with several different orbits on this set, because the determinant gives an invariant valued in $(\mathbf{Z}/N\mathbf{Z})^*/\{\pm 1\}$.

There's a similar situation for PGL₂. In particular, although there's a map $[SL_2] \rightarrow [PGL_2]$ it is not a homeomorphism. It is very far from injective *and* very far from surjective.

4.5. **[G]** for $G = SO_q$. Let q be an integral quadratic form on \mathbb{Z}^r , nondegenerate over \mathbb{Q} . Let SO_q be the orthogonal gorup and set $K = \prod_p SO_q(\mathbb{Z}_p)$. Then at least

$$[SO_a]/K$$

has a classical description: it is naturally in bijection with the *genus* of the quadratic form q.

4.6. **[G]** for general **G**. Now let **G** be a general reductive group over **Q**, and $K \subset \mathbf{G}(\mathbf{A}_f)$ an open compact subgroup. In general, $[\mathbf{G}]/K$ will be a union of finitely many components ADDREF, each of the form $\Gamma \setminus \mathbf{G}(\mathbf{R})$:

$$[\mathbf{G}]/K \simeq \coprod_{i \in I} \Gamma_i \backslash \mathbf{G}(\mathbf{R}).$$

The set of components *I* is itself indexed by something like a class number.

Here Γ is an "arithmetic subgroup" of $\mathbf{G}(\mathbf{Q})$. What this means is the following: Although the notion $\mathbf{G}(\mathbf{Z})$ doesn't make sense *a priori*, there is a well-defined commensurability class of subgroups of $\mathbf{G}(\mathbf{Q})$ that contains $\mathbf{G}(\mathbf{Z})$ whenever it's defined. In other words, $\mathbf{G}(\mathbf{Z})$ is "well-defined up to finite index ambiguities." Precisely, a subgroup $\Gamma \leq \mathbf{G}(\mathbf{Q})$ is called an *arithmetic subgroup* if Γ is commensurable to $\rho^{-1}(\mathrm{SL}_M(\mathbf{Z}))$ where $\rho: \mathbf{G} \to \mathrm{SL}_M$ is an embedding. The notion of arithmetic subgroup is independent of ρ .

5. Holomorphic forms

We have as usual the space $S_k(\Gamma(N))$ of modular cusp forms of level k for the subgroup $\Gamma(N)$ and the Hecke operators T_p , for p not dividing N. A *Hecke eigenform* is simply an eigenfunction of all such T_p .

We will construct an injection:

Hecke eigenforms of level $1 \hookrightarrow \text{Automorphic representations for } SL_2$.

The map is given by $f \mapsto SL_2(\mathbf{A})\tilde{f}$, i.e. we first will lift f to a function \tilde{f} on $SL_2(\mathbf{Z}) \setminus SL_2(\mathbf{R})$ and thus also (via (2)) on $[SL_2]$, then take its $SL_2(\mathbf{A})$ -translates, and then take the closure of their span.

To formulate a more general statement it is better to work with GL_2 rather than SL_2 . (We shall see why in the course of proof.) We say two Hecke eigenforms f, g are *nearly equivalent* if

$$\lambda_p(f) = \lambda_p(g)$$
 for all but finitely many primes p .

More generally, similar methods give rise to an injection

Hecke eigenforms modulo near equivalence \hookrightarrow Automorphic representations for GL_2 .

The condition of being a Hecke eigenform and being holomorphic will translate into what we want.

5.1. **Lifting to** $\Gamma(N)\backslash SL_2(\mathbf{R})$. The first point is that we can lift any modular form from $\Gamma\backslash \mathbb{H}$ to $\Gamma\backslash SL_2(\mathbf{R})$. That is to say, even though a modular form is not quite a function on $\Gamma\backslash \mathbb{H}$ (it is a section of a line bundle) it can be lifted to a function on $\Gamma\backslash SL_2(\mathbf{R})$ (the line bundle becomes trivial after pullback).

For weight 2 this process is particularly geometric: We can identify $SL_2(\mathbf{R})/\{\pm 1\}$ to the unit tangent bundle of \mathbb{H} . Indeed, $SL_2(\mathbf{R})/\{\pm 1\}$ acts on \mathbb{H} and so also on its unit tangent bundle and the action on the tangent bundle is simply transitive.

Now a weight 2 form is the same thing as a differential 1-form (i.e. f(z)dz is Γ -invariant, i.e. a function on the tangent bundle). This gives an association

weight 2 forms
$$\longrightarrow$$
 functions on $\Gamma \backslash SL_2(\mathbf{R})$

Explicitly, to the weight 2 form f we make the function

(6)
$$\tilde{f}: \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (ci+d)^{-2} f(\frac{ai+b}{ci+d}).$$

and we obtain a corresponding lifting for weight m forms f, replacing $(ci+d)^{-2}$ by $(ci+d)^{-m}$.

Note that for $\kappa = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in SO(2)$ – this is just the stabilizer of $i \in \mathbb{H}$ in-

side $SL_2(\mathbf{R})$ – we have $\tilde{f}(g\kappa) = \tilde{f}(g)e^{im\theta}$, where m is the weight of f. Therefore, we can regard – via (3) – \tilde{f} as a function on $SL_2(\mathbf{A})$ that satisfies

$$(k_1, \kappa) \cdot \tilde{f} = \chi_m(k) \tilde{f},$$

for
$$k_1 \in SO(2) \subset SL_2(\mathbf{R})$$
 and $\kappa \in SL_2(\widehat{\mathbf{Z}})$; here $\chi_m : \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \mapsto e^{im\theta}$.

5.2. Hecke algebras and the abstract irreducibility criterion. For a finite group G acting on a complex vector space V, the group algebra $\mathbf{C}G$ acts on V via

$$A \in \mathbf{C}G : v \mapsto \sum_{g \in G} A_g \cdot g \, v.$$

Here the multiplication on the group algebra is given by $(AB)_g = \sum_x A_x B_{x^{-1}g}$.

There is an analogue in the context of a "nice" ADDREF topological group G and a unitary representation V. The algebra of continuous, compactly supported functions $C_c(G)$ acts on V, via the rule

$$F \in C_c(G) : v \mapsto \int_{g \in G} F(g) dg \cdot g v,$$

where dg is a Haar measure. In fact, what matters here is not the function F(g) but the measure F(g)dg, and we can actually regard this as an action of the space of compactly supported complex-valued *measures*; then there needs to be no choice of Haar measure

at all. For example, if $K \subset G$ is a compact subgroup, and 1_K is the Haar probability measure on K, then 1_K acts on V as a projection onto the K-invariant subspace.

Suppose that V is a unitary representation of a group G, and $K \subset G$ is a compact subgroup. Then the fixed vectors V^K carry an extra structure – an action of the Hecke algebra $\mathcal{H}(G,K)$, that is to say, the K-bi-invariant continuous functions on G, considered as a convolution algebra.

Lemma: If $f \in V^K$ is a $\mathcal{H}(G,K)$ -eigenfunction then $U := \overline{G \cdot f}$ is irreducible.

Proof. First of all, $U^K = \langle f \rangle$:

Suppose that $u \in U^K$. Pick $\varepsilon > 0$. We can find $u' = \sum a_i(g_i \cdot f)$, for some $a_i \in \mathbb{C}$, $g_i \in G$, where $||u - u'|| < \varepsilon$. Then also $||u - 1_K u'|| < \varepsilon$, but

$$1_K u' = \sum a_i (1_K g_i 1_K f) \in \langle f \rangle.$$

So u lies in the closure of $\langle f \rangle$, i.e. $u \in \langle f \rangle$. Now suppose that $U = U_1 \oplus U_2$. By what we just said, U_1^K or U_2^K must be trivial; without loss it is U_2^K , but then $f \in U_1^K$ so $\overline{Gf} \subset U_1$, and so $U = U_1$.

This proves, in fact, the more general statement: If V is topologically generated by V^K , then V is irreducible if and only if V^K is irreducible as a \mathcal{H} -representation.

5.3. There are many ways to generalize this. Brian's Lecture 2 will be a more comprehensive treatment of the nonarchimedean local analogue.

In particular, suppose that $\chi: K \to S^1$ is a character, and set $V^{(K,\chi)} = \{v \in V : kv = \chi(k)v\}$. Then, similarly, if $f \in V^{(K,\chi)}$ is an eigenfunction of $\mathcal{H}(G,K;\chi)$ – the algebra of compactly supported measures μ that satisfy

$$\delta_{k_1}\mu\delta_{k_2}=\chi(k_1)\chi(k_2)\mu,$$

where δ_a is the point measure at a – then in fact \overline{Gf} is irreducible.

5.4. According to §5.3, then, in order to verify that $\overline{\operatorname{SL}_2(\mathbf{A}).\tilde{f}}$ is an irreducible subrepresentation of $L^2([\operatorname{SL}_2])$ we need to compute the action of a certain "Hecke algebra" on \tilde{f} .

As we sketch, this Hecke algebra is (topologically) generated by:

- Hecke operators T_{p^2} , for p a prime; these are related to $\mathcal{H}(\mathrm{SL}_2(\mathbf{Q}_p),\mathrm{SL}_2(\mathbf{Z}_p);$
- Circle-averaging operators; these are related to $\mathcal{H}(SL_2(\mathbf{R}), (SO_2, \chi))$.

To be updated after Oct 22 lecture -AV.