AUTOMORPHIC FORMS ON QUATERNION ALGEBRAS

ANDREW SNOWDEN

Let F be a totally real number field and let D be a quaternion algebra over F. Define an algebraic group G over F by $G(A) = (A \otimes_F D)^\times$ for an F-algebra A. It is easy to see that if F'/F splits D then $G_{F'}$ is isomorphic to $\mathrm{GL}(2)$. Thus G is a reductive algebraic group. We therefore have a theory of automorphic forms and representations for G. We will look at the basics of this theory in these notes.

To begin with, let us examine the spaces on which automorphic forms are functions. Let f be an automorphic form on $G(\mathbf{A}_F)$. By definition, f is a function $G(\mathbf{A}_F) \to \mathbf{C}$ subject to the following:

- f is left invariant under $G(F)$.
- f is invariant under a compact open subgroup U of $G(\mathbf{A}_{F,f})$.
- f is finite under translations by a maximal compact subgroup K of $G(F \otimes \mathbf{R})$.
- f is finite under the center of the universal enveloping algebra of $G(F \otimes \mathbf{R})$.
- f satisfies certain continuity and growth conditions.

For simplicity, let us consider the case where f is invariant under the center Z of $G(\mathbf{A}_F)$ and transforms under K by a one dimensional representation σ, i.e., $f(gk) = \sigma(k)f(g)$ holds. Then f defines a section of a line bundle determined by σ on the space

$$X(U) = G(F)\backslash G(\mathbf{A}_F)/ZKU.$$

Our first task is to describe this space.

The most important thing to initially consider about $X(U)$ is the contributions of the infinite places. At an infinite place v the division algebra D has two possibles behaviors: it can either split or not. In either case, $D_v^\times$ is a four dimensional real Lie group. If D_v is split then G_v looks like $\mathrm{GL}_2(\mathbf{R})$ and so its maximal compact is the two dimensional orthogonal group. We thus find that G_v/K_vZ_v is a copy of the upper half plane. In particular, it is a one dimensional complex manifold. If D_v is non-split then G_v is the multiplicative group in the Hamilton quaternions. This group is an extension of the rank 2 unitary group by $\mathbf{R}_+$. Thus $G_v = Z_vK_v$ and so the quotient G_v/Z_vK_v is a point. We therefore find

$$G(F \otimes \mathbf{R})/Z_\infty K = \mathfrak{h}^n$$

where n is the number of infinite places at which D is split. This computation is significant for two reasons. First, we see that the quotient is canonically a complex manifold, so we can make sense of holomorphic functions on it. And secondly, if $n = 0$, that is, if D is non-split at all infinite places, then this space is just a point.

Consider the case $n > 0$. Since $G(F \otimes \mathbf{R})$ is non-compact, strong approximation gives

$$G(\mathbf{A}_F) = G(F)G(F \otimes \mathbf{R})U$$

and so the usual computation show that

$$X(U) = \Gamma(U)\backslash \mathfrak{h}^n$$

where $\Gamma(U)$ is the arithmetic group obtained from intersecting $G(F)$ and U. (This assumes something about U: the norm map $U \to \prod \mathbf{G}_m(\mathscr{O}_{F,v})$ is surjective. For a general U, $X(U)$ will not be connected but will have finitely many connected components, each of the above form.) Thus $X(U)$ is a complex manifold obtained as the quotient of n copies of the upper half plane by the action of a totally discontinuous subgroup. When $n = 1$ (and $F \neq \mathbf{Q}$) these spaces are called *Shimura curves*. In contrast to the modular curves, they are compact — no cusps need to be added to obtain a compact space.

Now consider the case $n = 0$. Strong approprixmation no longer applies to $G(F \otimes \mathbf{R})$ since this group is compact. However, we do not really need to use strong approximation. Since KZ contains all of $G(F \otimes \mathbf{R})$

Date: January 28, 2010.

we can completely ignore the infinite places. We find

$$X(U) = G(F)\backslash G(\mathbf{A}_{F,f})/Z_f U.$$

Since U is a compact open subgroup of $G(\mathbf{A}_{F,f})$, this quotient is discrete. In fact, it is a finite set; this follows from it being discrete and having finite volume. Thus in this case, automorphic forms on D are simply functions on a finite set. There are therefore no continuity or analytic conditions placed on the forms — they are just functions on a finite set!

We did not consider the most general possible set-up. One does not need to assume Z-invariance: one can allow f to transform under Z by a character. Also, the representation σ of K, which plays the role of the weight, does not have to be one dimensional. (One must then take f to be vector valued.) The rank two unitary group has irreducible representations of all dimensions, so when D is non-split at infinite places one might want to consider such representations of K.

We now change directions and consider automorphic representations of $G(\mathbf{A}_F)$. Such a representation decomposes as a tensor product $\otimes\pi_v$ over the places of F, where π_v is an irreducible admissible representation of G_v. For almost all places, D_v is split, and so G_v is isomorphic to $\mathrm{GL}_2(F_v)$. The local invariants (such as conductor, L-series and ϵ-factors) at these places are defined as usual. We now consider a place v at which D_v is non-split. The group G_v is compact modulo its center. This implies that the representation π_v is finite dimensional. The maximal compact subgroup K_v of G_v has a natural filtration $K_v^{(n)}$ obtained by looking at the group of elements congruent to 1 modulo powers of the maximal ideal. There is a unique minimal n such that π_v contains $K_v^{(n)}$ in its kernel. The number $n+1$ is called the *conductor* of π_v (or perhaps the exponent of the conductor). The prime power $\mathfrak{p}_v^{n+1}$ is the local contribution of π_v to the level. Note that this exponent is never 0 — even when π_v contains the full maximal compact subgroup K_v in its kernel the conductor is 1. The reason for this will be evident later — suffice it to say for now that the Galois representations coming from modular forms on a quaternion algebra are always ramified where the quaternion algebra is, so these places should appear in the conductor.

We can additionally attach local L-functions and ϵ-factors to π_v. This goes quite similarly to the GL(1) case (Tate's thesis). For a Schwartz function ϕ on D_v one considers the integral

$$Z(s, \phi, \pi_v) = \int_{D_v} \phi(x) \operatorname{tr} \pi_v(x) |x|^{s+\cdots} dx.$$

The elipses in the exponent is a normalizing factor, which is not important for the present discussion. One then finds that there is a unique Euler factor $L(s, \pi_v)$ such that the quotient

$$\frac{Z(s, \phi, \pi_v)}{L(s, \pi_v)}$$

is an entire function of s and for some choice of ϕ is equal to 1. Furthermore, there is a functional equation

$$\frac{Z(1-s, \hat{\phi}, \pi_v^\vee)}{L(1-s, \pi_v^\vee)} = \epsilon(s, \pi_v, \psi_v) \frac{Z(s, \phi, \pi_v)}{L(s, \pi_v)}.$$

Here ψ_v is a non-trivial additive character of F_v, $\hat{\phi}$ denotes the Fourier transform of ϕ with respect to ψ_v and $\pi_v^\vee$ denotes the contragredient of π_v. The factor $\epsilon(s, \pi_v, \psi_v)$ is of the form $s \mapsto ab^s$. The base b of this exponential is equal (or almost equal) to the conductor of π_v.